

geron

U.S. UTILITY PATENT APPLICATION

Title: SELECTIVE ANTIBODY TARGETING OF
UNDIFFERENTIATED STEM CELLS

**Filing
Information:** Date: November 26, 2001
Express Mail No. EF 279099883 US
Attorney Docket: 096-004

Inventors: Jim McWhir
Midlothian, Scotland, U.K.

Joseph D. Gold
San Francisco, CA

J. Michael Schiff
Menlo Park, CA

Correspondence: David J. Earp
Registration No. 41,401

Geron Corporation
230 Constitution Drive
Menlo Park, CA 94025
Phone: (650) 473-7700
Fax: (650) 473-8654

USPTO Customer No. 022869

Priority Data

This application claims priority to U.S. Patent Applications 60/253,357;
60/253,443; and 60/253,395, all filed November 27, 2001, pending.

SELECTIVE ANTIBODY TARGETING OF UNDIFFERENTIATED STEM CELLS

5

TECHNICAL FIELD

This invention relates generally to the field of cell biology of stem cells, embryonic cells, and the molecular biology of promoter controlled viral vectors. More specifically, it describes a technology for removing undifferentiated cells from populations derived from pluripotent stem cells using selectively expressed lytic vectors.

10

REFERENCE TO RELATED APPLICATIONS

This application claims priority to U.S. Patent Applications 60/253,357; 60/253,443; and 60/253,395, all filed November 27, 2001, pending. The priority documents are hereby incorporated herein by reference in their entirety.

15

BACKGROUND

Precursor cells have become a central interest in medical research. Many tissues in the body have a back-up reservoir of precursors that can replace cells that are senescent or damaged by injury or disease. Considerable effort has been made recently to isolate precursors of a number of different tissues for use in regenerative medicine.

20

U.S. Patent 5,750,397 (Tsukamoto et al., Systemix) reports isolation and growth of human hematopoietic stem cells which are Thy-1+, CD34+, and capable of differentiation into lymphoid, erythroid, and myelomonocytic lineages. U.S. Patent 5,736,396 (Bruder et al.) reports methods for lineage-directed differentiation of isolated human mesenchymal stem cells, using an appropriate bioactive factor. The derived cells can then be introduced into a host for mesenchymal tissue regeneration or repair.

25

U.S. Patent 5,716,411 (Orgill et al.) proposes regenerating skin at the site of a burn or wound, using an epithelial autograft. U.S. Patent 5,766,948 (F. Gage) reports a method for producing neuroblasts from animal brain tissue. U.S. Patent 5,672,499 (Anderson et al.) reports obtaining neural crest stem cells from embryonic tissue. U.S. Patent 5,851,832 (Weiss et al., Neurospheres) reports isolation of putative neural stem cells from 8-12 week old human fetuses. U.S. Patent 5,968,829 (M. Carpenter) reports human neural stem cells derived from primary central nervous system tissue.

30

U.S. Patent 5,082,670 (F. Gage) reports a method for grafting genetically modified cells to treat defects, disease or damage of the central nervous system. Auerbach et al. (Eur. J. Neurosci. 12:1696, 2000) report that multipotential CNS cells implanted into animal brains form electrically active and functionally connected neurons. Brustle et al. (Science 285:754, 1999) report that precursor cells derived from embryonic stem cells interact with host neurons and efficiently myelinate axons in the brain and spinal cord.

35

Considerable interest has been generated by the development of embryonic stem cells, which are thought to have the potential to differentiate into many cell types. Early work on embryonic stem cells was done in mice. Mouse stem cells can be isolated from both early embryonic cells and germinal tissue. Desirable characteristics of

40

pluripotent stem cells are that they be capable of proliferation in vitro in an undifferentiated state, retain a normal karyotype, and retain the potential to differentiate to derivatives of all three embryonic germ layers (endoderm, mesoderm, and ectoderm).

Development of human pluripotent stem cell preparations is considerably less advanced than work with mouse cells. Thomson et al. propagated pluripotent stem cells from lower primates (U.S. Patent 5,843,780; Proc. Natl. Acad. Sci. USA 92:7844, 1995), and then from humans (Science 282:114, 1998). Gearhart and coworkers derived human embryonic germ (hEG) cell lines from fetal gonadal tissue (Shamblott et al., Proc. Natl. Acad. Sci. USA 95:13726, 1998; and U.S. Patent 6,090,622).

Both hES and hEG cells have the long-sought characteristics of pluripotent stem cells: they are capable of being grown in vitro without differentiating, they have a normal karyotype, and they remain capable of producing a number of different cell types. Clonally derived human embryonic stem cell lines maintain pluripotency and proliferative potential for prolonged periods in culture (Amit et al., Dev. Biol. 227:271, 2000). These cells hold considerable promise for use in human therapy, acting as a reservoir for regeneration of almost any tissue compromised by genetic abnormality, trauma, or a disease condition.

International Patent Publication WO 99/20741 (Geron Corp.) refers to methods and materials for growing primate-derived primordial stem cells. In one embodiment, a cell culture medium is provided for growing primate-derived primordial stem cells in a substantially undifferentiated state, having a low osmotic pressure and low endotoxin levels. The basic medium is combined with a nutrient serum effective to support the growth of primate-derived primordial stem cells and a substrate of feeder cells or an extracellular matrix component derived from feeder cells. The medium can further include non-essential amino acids, an anti-oxidant, and growth factors that are either nucleosides or a pyruvate salt.

A significant challenge to the use of stem cells for therapy is to control growth and differentiation into the particular type of tissue required for treatment of each patient.

U.S. Patent 4,959,313 (M. Taketo, Jackson Labs) provides a particular enhancer sequence that causes expression of a flanking exogenous or recombinant gene from a promoter accompanying the gene that does not normally cause expression in undifferentiated cells. U.S. Patent 5,639,618 (D.A. Gay, Plurion Inc.) proposes a method for isolating a lineage specific stem cell in vitro, in which a pluripotent embryonic stem cell is transfected with a construct in which a lineage-specific genetic element is operably linked to a reporter gene, culturing the cell under conditions where the cell differentiates, and then separation of cells expressing the reporter are separated from other cells.

U.S. Patent 6,087,168 (Levesque et. al., Cedars Sinai Med. Ctr.) is directed to transdifferentiating epidermal cells into viable neurons useful for both cell therapy and gene therapy. Skin cells are transfected with a neurogenic transcription factor, and cultured in a medium containing an antisense oligonucleotide corresponding to a negative regulator of neuronal differentiation.

International Patent Publication WO 97/32025 (McIvor et al., U. Minnesota) proposes a method for engrafting drug-resistant hematopoietic stem cells. The cells in the graft are augmented by a drug resistance gene (such as methotrexate resistant dihydrofolate reductase), under control of a promoter functional in stem cells. The cells are administered into a mammal, which is then treated with the drug to increase engraftment of transgenic cells relative to nontransgenic cells.

International Patent Publication WO 98/39427 (Stein et al., U. Massachusetts) refers to methods for expressing exogenous genes in differentiated cells such as skeletal tissue. Stem cells (e.g., from bone marrow) are contacted with a nucleic acid in which the gene is linked to an element that controls expression in differentiated

cells. Exemplary is the rat osteocalcin promoter. International Patent Publication WO 99/10535 (Liu et al., Yale U.) proposes a process for studying changes in gene expression in stem cells. A gene expression profile of a stem cell population is prepared, and then compared a gene expression profile of differentiated cells

International Patent Publication WO 99/19469 (Braetscher et al., Biotransplant) refers to a method for growing pluripotent embryonic stem cells from the pig. A selectable marker gene is inserted into the cells to be regulated by a control or promoter sequence in the ES cells, exemplified by the porcine OCT-4 promoter.

International Patent Publication WO 00/15764 (Smith et al., U. Edinburgh) refers to propagation and derivation of embryonic stem cells. The cells are cultured in the presence of a compound that selectively inhibits propagation or survival of cells other than ES cells by inhibiting a signaling pathway essential for the differentiated cells to propagate. Exemplary are compounds that inhibit SHP-2, MEK, or the ras/MAPK cascade.

Klug et al. (J. Clin. Invest. 98:216, 1996) propose a strategy for genetically selecting cardiomyocytes from differentiating mouse embryonic stem cells. A fusion gene consisting of the α -cardiac myosin heavy chain promoter and a cDNA encoding aminoglycoside phosphotransferase was stably transfected into the ES cells. The resulting lines were differentiated in vitro and selected using G418. The selected cardiomyocyte cultures were reported to be highly differentiated. When engrafted back into mice, ES-derived cardiomyocyte grafts were detectable as long as 7 weeks after implantation.

Schuldiner et al. (Proc. Natl. Acad. Sci. USA 97:11307, 2000) report the effects of eight growth factors on the differentiation of cells from human embryonic stem cells. After initiating differentiation through embryoid body formation, the cells were cultured in the presence of bFGF, TGF- β 1, activin-A, BMP-4, HGF, EGF, β NGF, or retinoic acid. Each growth factor had a unique effect on the differentiation pathway, but none of the growth factors directed differentiation exclusively to one cell type.

There is a need for new approaches to generate populations of differentiated cells suitable for human administration.

SUMMARY OF THE INVENTION

This invention provides a system for depleting relatively undifferentiated cells from a heterogeneous cell population, such as may be obtained by differentiation of stem cells. The population is treated with a vector that puts an effector gene under control of a gene element that allows the gene to be expressed at a higher level in the undifferentiated subpopulation. Exemplary effector genes are glycosyltransferases, rendering undifferentiated cells separable using specific antibody, or susceptible to lysis by antibody plus complement. This produces a population relatively enriched for mature cells, and suitable for use in regenerative medicine.

One embodiment of this invention is a method of producing differentiated cells. A cell population comprising undifferentiated stem cells that contain a nucleic acid molecule comprising the structure P-X is treated so as to cause at least some undifferentiated cells in the population to differentiate. X is nucleic acid sequence that causes expression of a cell surface antigen under control of transcriptional control element P, which has the effect of causing the surface antigen to be preferentially expressed in undifferentiated cells. The connecting line in P-X indicates that the genetic elements are operatively linked, whether or not they are adjacent in the nucleic acid molecule.

After the cell population is differentiated, relatively undifferentiated cells can be depleted from the population by combining the cells with ligand specific for the antigen. In this context, the term "ligand" refers to any biological molecule (typically a protein) that binds the antigen with a specificity to discriminate the antigen from

other molecules on the cell surface or on other cells in the population. Suitable ligands include specific monoclonal or polyclonal antibody and lectins. Depletion can be effected, for example, by combining the cells with ligand specific for the antigen, and separating cells that have not bound the ligand. Where the specific ligand is an antibody, it can be combined with the cells in culture, or the cells can be placed in a subject having circulating natural antibody, or antibody that has been induced by active or passive immunization. The undifferentiated cells can be removed, for example, by an affinity separation technique, cell sorting, or by complement-mediated lysis in culture or in situ.

Another embodiment of the invention is a method for depleting undifferentiated stem cells from a cell population. Stem cells in the population are genetically altered so that they contain a nucleic acid molecule comprising the structure P-X as already described. Undifferentiated cells are then depleted from the population using antibody specific for the cell surface antigen. The cell population can be genetically altered when it is still predominantly undifferentiated (before being caused to differentiate), or when it already predominantly comprises differentiated cells.

In certain embodiments, X encodes a transmembrane protein which itself acts as the cell surface antigen. In other embodiments, X encodes an enzyme that in turn causes the antigen to be expressed on the cell surface. Exemplary are glycosyltransferase enzymes, particularly $\alpha(1,3)$ galactosyltransferase, which causes expression of the Gal $\alpha(1,3)$ Gal xenoantigen, and ABO blood group transferases, which cause expression of the ABO histo blood group alloantigens. In order to enhance dominance of the desired phenotype, the glycosyltransferase encoding region can incorporate a heterologous membrane anchoring segment or cytoplasmic domain (either at the N- or C-terminus) that optimizes positioning of the enzyme within the Golgi apparatus.

In certain embodiments, P-X is an introduced heterologous molecule, meaning that the cell or its ancestors was genetically altered with a vector comprising P-X. In other embodiments the cell or its ancestors was genetically altered with a vector to place X under control of an endogenous transcriptional control element. Following transfection, X can be either transiently expressed in undifferentiated cells in the population, or P-X can be inheritable and expressed in undifferentiated progeny. Non-limiting examples for P include the OCT-4 promoter, and the promoter of telomerase reverse transcriptase (TERT). The cells can also contain a drug resistance gene Y under control of P.

A further embodiment of the invention is a stem cell genetically altered to express a carbohydrate antigen not normally expressed by the cell, possibly an antigen recognized by a naturally occurring antibody. As an example, the cell can be genetically altered with a glycosyl transferase, such as an $\alpha(1,3)$ galactosyltransferase, or an ABO blood group transferase. Expression of the carbohydrate antigen can be controlled by a transcriptional control element specific for undifferentiated cells.

The reagents and techniques of this invention can be brought to bear on cell populations containing many different types of stem cells, as described below. They are especially suited for application to primate pluripotent stem cells, such as human embryonic stem cells.

Other embodiments of the invention will be apparent from the description that follows.

BRIEF DESCRIPTION OF THE DRAWINGS

Figur 1 provides an analysis of OCT-4 and hTERT expression in hES cells cultured with feeder cells (mEF) or extracellular matrix (Matrigel® or laminin) with regular medium (RM) or conditioned medium (CM). The

upper panel is a copy of a gel showing OCT-4 and hTERT expression at the mRNA level by RT-PCR. The lower panel is a bar graph comparing the level of expression for cells grown on different substrates, expressed as the ratio of OCT-4 or hTERT to the 18s standard. hES cells grown on Laminin and Matrigel® in conditioned medium have similar expression patterns to those of cells grown on a feeder layer.

Figure 2 is a half-tone reproduction of a gel showing telomerase activity measured in cultured hES cells by TRAP activity assay. All the culture conditions showed positive telomerase activity after 40 days in feeder-free culture.

Figure 3 is a half-tone reproduction showing expression of the GFP reporter gene in hES cells transduced with retrovirus and then differentiated. hES cells were transferred to suspension culture to form embryoid bodies, cultured for a further 4 days, replated onto gelatin-coated slides and cultured for a week, and then fixed and photographed under fluorescence for GFP expression. Left panels show bright-field illumination; right panels show fluorescence due to GFP expression.

Figure 4 shows the results of a study in which hES cells were transiently genetically altered in feeder-free culture by lipofection. Panel A is a half-tone reproduction of a light micrograph showing morphology of hES cells on laminin after they have been transfected. Panel B is a half-tone reproduction of a fluorescence micrograph showing GFP expression in the same colony. Panel C is a bar graph showing percentage of cells expressing GFP under various conditions.

Figure 5 is a map of TPAC vector designated pGRN376. This is an adenovirus vector of 7185 bp comprising the herpes simplex thymidine kinase (*tk*) gene under control of a promoter taken from the upstream sequence of the human gene for telomerase reverse transcriptase (hTERT). Expression of *tk* is promoted in cells expressing hTERT, such as undifferentiated embryonic stem cells.

Figure 6 is a two-panel line graph, showing the effect of the TPAC thymidine kinase vector on undifferentiated hES cells. 48 h after replating, the cells were transduced with TPAC vector at an MOI of 30 or 100, or mock transduced (no vector added). Four h later, the cells were exchanged into fresh medium containing the prodrug ganciclovir (GCV). By day 3, wells treated with TPAC vector + GCV contained 8% as many cells as the control wells.

Figure 7 is a bar graph showing titration of GCV in TPAC vector treated hES cells. 4 h after transduction with the vector, fresh medium was added containing GCV at the concentration shown. ~20 μ M GCV was optimal under the conditions tested.

Figure 8 is a two-panel bar graph showing titration of GCV on TPAC vector transduced and mock-transduced hES cells from two different lines. Both lines are sensitive to GCV after treatment with the TPAC vector.

Figure 9 shows the effect of TPAC + GCV treatment on mixed cell populations obtained from differentiation of hES cells. The cells were fed daily with conditioned medium to maintain the undifferentiated state, or with either 500 nM retinoic acid or 0.5% DMSO, to induce differentiation into committed cells of mixed phenotype. 7 days later, they were infected with the TPAC vector at an MOI of 30, plus 20 μ M GCV.

The Upper Panel is a bar graph showing the number of cells surviving in culture. Treatment with TPAC + GCV eliminated cells cultured under each condition. In each instance, culture of the surviving cells produced populations that appeared highly differentiated and substantially free of undifferentiated morphology. The Lower Panel is a half-tone reproduction of a gel showing RT-PCR analysis of the surviving cells. Those cells cultured with conditioned medium (mEF-CM) or DMSO had no detectable OCT-4 expression, while 2 out of 4 samples treated with retinoic acid (RA) showed amplification products consistent with very low levels of OCT-4 expression.

Figure 10 is a schematic depiction of targeting strategy to place an $\alpha(1,3)$ galactosyltransferase encoding sequence under control of the endogenous hTERT promoter on one allele (Example 13). In the two-step approach (upper panel), the endogenous hTERT gene is targeted with a promoterless vector comprising the *neo* gene, and selected for G418 resistance. The *neo* sequence is then replaced with $\alpha(1,3)$ galactosyltransferase ($\alpha 1,3$ GT) using *cre* recombinase. In the one step approach (lower panel), *neo* is introduced with an internal ribosomal entry site 3' to the $\alpha 1,3$ GT coding region. In this instance, the $\alpha 1,3$ GT is truncated before the polyadenylation signal and is transcribed directly from the hTERT promoter. A bicistronic message is produced from which both proteins are translated.

DETAILED DESCRIPTION OF THE INVENTION

Stem cells of various kinds have become an extremely attractive modality in regenerative medicine. They can be proliferated in culture, and then differentiated in vitro or in situ into the cell types needed for therapy. Recently, it has been demonstrated that human embryonic stem cells continuously express a high level of telomerase, enabling them to maintain telomere length and grow almost indefinitely in culture.

So far, efforts to differentiate stem cells have been directed primarily towards identifying culture conditions that promote outgrowth of a cell population with phenotypic features of a tissue type desirable for regenerative medicine. Schuldiner et al. (*supra*) report the effects of growth factors on the differentiation of human embryonic stem cells. In U.S. Patent 5,639,613, stem cells are transfected with a lineage-specific gene that is operably linked to a reporter gene, which is then used to select for cells expressing the reporter. In WO 97/32025, hematopoietic stem cells are augmented by a drug resistance gene, and then engrafted into a subject. The cells are administered into a mammal, which is then treated with the drug to increase engraftment of transgenic cells. Klug et al. (*supra*) used a construct in which the α -cardiac myosin heavy chain promoter controlled expression of aminoglycoside phosphotransferase. Transfected differentiated cells were selected using G418, which produced lines of cardiomyocyte like cells. This is a positive selection strategy that uses gene expression patterns of the desired tissue type to allow preferential survival of differentiated tissue.

It is a hypothesis of this invention that positively selecting for differentiated cells produces populations that are suboptimal for use in human therapy. Any undifferentiated cells in the population may impair engraftment or function of the cells in vivo. Undifferentiated cells may also increase the possibility of a malignancy or other tumor forming at the site of the therapeutic implant, or by migration of transplanted cells.

This invention is directed towards a strategy in which undifferentiated cells remaining in a differentiated cell population are depleted. This is effected by genetically altering the cells, so that a gene that is lethal to a cell in which it is expressed, or renders it susceptible to a lethal effect of an external agent, is placed under transcriptional control of a genetic element that causes it to be expressed preferentially in any undifferentiated cells in the population. This is a negative selection strategy, designed to minimize the proportion of undifferentiated cells. It is possible to combine this technique with positive selection techniques of various kinds, in order to obtain relatively pure populations of the desired tissue type that are essentially free of undifferentiated cells.

In certain embodiments of the invention, the cell population is transfected with a genetic construct in which a promoter specific for undifferentiated cells drives a glycosyl transferase, which in turn synthesizes a new surface antigen. For example, $\alpha(1,3)$ galactosyltransferase ($\alpha 1,3$ GT) can be used to express the Gal $\alpha(1,3)$ Gal epitope on the cell surface. After culturing the cells under conditions where the antigen can be formed, the cells

are separated using specific antibody, or treated with specific antibody and complement to deplete undifferentiated cells from the population. A number of glycosyl transferases are suitable for this purpose, particularly those that synthesize a xenoantigen or alloantigen against which humans have naturally occurring antibody.

To validate the negative selection strategy, human embryonic stem (hES) cells have been transduced with an adenovirus vector (TPAC) in which a herpes virus thymidine kinase gene was placed under control of a promoter sequence for human telomerase reverse transcriptase (hTERT). hES cells constitutively express hTERT, but this ability is lost upon differentiation. Example 10 (Figures 6-8) show that transduction of hES cells with TPAC vector renders undifferentiated cells susceptible to lethality by the prodrug ganciclovir, a substrate for thymidine kinase, at a concentration of ~ 20 μ M. Example 11 (Figure 9) shows that when hES cells are transduced with TPAC vector and then differentiated with DMSO, there are no surviving cells with detectable OCT-4 expression (a phenotype of undifferentiated cells).

The techniques of this invention are designed in part to provide cell populations with improved characteristics for human therapy. After depleting undifferentiated cells, the differentiated population is expected to possess better functional and engraftment characteristics, and have reduced risk of creating unwanted tissue architecture and malignancies in the treated subject. In addition, cell populations depleted of undifferentiated cells are more homogeneous, which provides a distinct advantage for non-therapeutic applications, such as producing antibody, cDNA libraries, and screening drug candidates.

A particular advantage of using a glycosyl transferase as the effector sequence is that the system provides *ongoing surveillance* after the cells are used for tissue regeneration. If undifferentiated cells reappear in the transplanted tissue (either through dedifferentiation or outgrowth of a preexisting subpopulation), the specific promoter will prompt synthesis of the glycosyltransferase — leading to expression of the antigen, followed by lysis of the undifferentiated cells in situ.

Another advantage is that the stem cells can be genetically altered with the glycosyl transferase in advance, and passaged or stored until use. The stem cell line can then be differentiated when needed, and depleted of undifferentiated cells by separating or lysing cells expressing the carbohydrate determinant synthesized by the transferase.

Definitions

Prototype "primate Pluripotent Stem cells" (pPS cells) are pluripotent cells derived from pre-embryonic, embryonic, or fetal tissue at any time after fertilization, and have the characteristic of being capable under appropriate conditions of producing progeny of several different cell types that are derivatives of all of the three germinal layers (endoderm, mesoderm, and ectoderm), according to a standard art-accepted test, such as the ability to form a teratoma in 8-12 week old SCID mice.

Included in the definition of pPS cells are embryonic cells of various types, exemplified by human embryonic stem (hES) cells, described by Thomson et al. (Science 282:1145, 1998); embryonic stem cells from other primates, such as Rhesus stem cells (Thomson et al., Proc. Natl. Acad. Sci. USA 92:7844, 1995), marmoset stem cells (Thomson et al., Biol. Reprod. 55:254, 1996) and human embryonic germ (hEG) cells (Shamblott et al., Proc. Natl. Acad. Sci. USA 95:13726, 1998). Other types of pluripotent cells are also included in the term. Any cells of primate origin that are capable of producing progeny that are derivatives of all three germinal layers are included, regardless of whether they were derived from embryonic tissue, fetal tissue, or other sources. This invention relates to pPS cells that are not derived from a malignant source. It is desirable (but not always necessary) that the cells be karyotypically normal.

pPS cell cultures are described as “undifferentiated” when a substantial proportion of stem cells and their derivatives in the population display morphological characteristics of undifferentiated cells, clearly distinguishing them from differentiated cells of embryo or adult origin. Undifferentiated pPS cells are easily recognized by those skilled in the art, and typically appear in the two dimensions of a microscopic view in colonies of cells with high nuclear/cytoplasmic ratios and prominent nucleoli. It is understood that colonies of undifferentiated cells within the population will often be surrounded by neighboring cells that are differentiated. Nevertheless, the undifferentiated colonies persist when the population is cultured or passaged under appropriate conditions, and individual undifferentiated cells constitute a substantial proportion of the cell population. Cultures that are substantially undifferentiated contain at least 20% undifferentiated pPS cells, and may contain at least 40%, 60%, or 80% in order of increasing preference. Whenever a culture or cell population is referred to in this disclosure as proliferating “without differentiation”, what is meant is that after proliferation, the composition is substantially undifferentiated according to the preceding definition.

“Feeder cells” or “feeders” are terms used to describe cells of one type that are co-cultured with cells of another type, to provide an environment in which the cells of the second type can grow. The feeder cells are optionally from a different species as the cells they are supporting. For example, certain types of pPS cells can be supported by primary mouse embryonic fibroblasts, immortalized mouse embryonic fibroblasts, or human fibroblast-like cells differentiated from hES cells, as described later in this disclosure. pPS cell populations are said to be “essentially free” of feeder cells if the cells have been grown through at least one round after splitting in which fresh feeder cells are not added to support the growth of the pPS. Cultures essentially free of feeder cells contain less than about 5% feeder cells. Whenever a culture or cell population is referred to in this disclosure as “feeder-free”, what is meant is that the composition is essentially free of feeder cells according to the preceding definition, subject only to further constraints explicitly required.

The term “embryoid bodies” is a term of art synonymous with “aggregate bodies”. The terms refer to aggregates of differentiated and undifferentiated cells that appear when pPS cells overgrow in monolayer cultures, or are maintained in suspension cultures. Embryoid bodies are a mixture of different cell types, typically from several germ layers, distinguishable by morphological criteria.

The terms “committed precursor cells”, “lineage restricted precursor cells” and “restricted developmental lineage cells” all refer to cells that are capable of proliferating and differentiating into several different cell types, with a range that is typically more limited than pluripotent stem cells of embryonic origin capable of giving rise to progeny of all three germ layers. Non-limiting examples of committed precursor cells include hematopoietic cells, which are pluripotent for various blood cells; hepatocyte progenitors, which are pluripotent for bile duct epithelial cells and hepatocytes; and mesenchymal stem cells. Another example is neural restricted cells, which can generate glial cell precursors that progress to oligodendrocytes and astrocytes, and neuronal precursors that progress to neurons.

For the purposes of this description, the term “stem cell” can refer to either a pluripotent stem cell, or a committed precursor cell, both as defined above. Minimally, a stem cell has the ability to proliferate and form cells of more than one different phenotype, and is also capable of self renewal — either as part of the same culture, or when cultured under different conditions. Embryonic stem cells can be identified as positive for the enzyme telomerase.

As used in this disclosure, “differentiated” and “undifferentiated” are relative terms depending on the context in which they are used. Specifically, in reference to a particular type of self-renewing stem cell, the term “undifferentiated” refers back to the same self-renewing stem cell, whereas the term “differentiated” refers to one or

more of the relatively mature phenotypes the stem cell can generate — as discernable by morphological criteria, antigenic markers, and gene transcripts they produce. Undifferentiated pPS cells have the ability to differentiate into all three germ layers. The cells differentiated from them do not, and can readily be recognized by one skilled in the art by morphological criteria.

The terms “polynucleotide” and “nucleic acid molecule” refer to a polymer of nucleotides of any length. Included are genes and gene fragments, mRNA, tRNA, rRNA, ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA and RNA, nucleic acid probes, and primers. As used in this disclosure, the term polynucleotides refer interchangeably to double- and single-stranded molecules. Unless otherwise specified or required, any embodiment of the invention that is a polynucleotide encompasses both a double-stranded form, and each of the two complementary single-stranded forms known or predicted to make up the double-stranded form. Included are nucleic acid analogs such as phosphoramidates and thiophosphoramidates.

A cell is said to be “genetically altered”, “transfected”, or “genetically transformed” when a polynucleotide has been transferred into the cell by any suitable means of artificial manipulation, or where the cell is a progeny of the originally altered cell that has inherited the polynucleotide. The polynucleotide will often comprise a transcribable sequence encoding a protein of interest, which enables the cell to express the protein at an elevated level. The genetic alteration is said to be “inheritable” if progeny of the altered cell have the same alteration.

A “control element” or “control sequence” is a nucleotide sequence involved in an interaction of molecules that contributes to the functional regulation of a polynucleotide, such as replication, duplication, transcription, splicing, translation, or degradation of the polynucleotide. Transcriptional control elements include promoters, enhancers, and repressors.

Particular gene sequences referred to as promoters, like the “TERT promoter”, or the “OCT-4 promoter”, are polynucleotide sequences derived from the gene referred to that promote transcription of an operatively linked gene expression product. It is recognized that various portions of the upstream and intron untranslated gene sequence may in some instances contribute to promoter activity, and that all or any subset of these portions may be present in the genetically engineered construct referred to. The promoter may be based on the gene sequence of any species having the gene, unless explicitly restricted, and may incorporate any additions, substitutions or deletions desirable, as long as the ability to promote transcription in the target tissue. Genetic constructs designed for treatment of humans typically comprise a segment that is at least 90% identical to a promoter sequence of a human gene. A particular sequence can be tested for activity and specificity, for example, by operatively linking to a reporter gene (Example 9).

Genetic elements are said to be “operatively linked” if they are in a structural relationship permitting them to operate in a manner according to their expected function. For instance, if a promoter helps initiate transcription of the coding sequence, the coding sequence can be referred to as operatively linked to (or under control of) the promoter. There may be intervening sequence between the promoter and coding region so long as this functional relationship is maintained.

In the context of encoding sequences, promoters, and other genetic elements, the term “heterologous” indicates that the element is derived from a genotypically distinct entity from that of the rest of the entity to which it is being compared. For example, a promoter or gene introduced by genetic engineering techniques into an animal of a different species is said to be a heterologous polynucleotide. An “endogenous” genetic element is an element that is in the same place in the chromosome where it occurs in nature, although other elements may be artificially introduced into a neighboring position.

The terms "polypeptide", "peptide" and "protein" are used interchangeably in this disclosure to refer to polymers of amino acids of any length. The polymer may comprise modified amino acids, it may be linear or branched, and it may be interrupted by non-amino acids.

The term "antibody" as used in this disclosure refers to both polyclonal and monoclonal antibody. The ambit of the term deliberately encompasses not only intact immunoglobulin molecules, but also such fragments and genetically engineered derivatives of immunoglobulin molecules as may be prepared by techniques known in the art, and which retains the binding specificity of the antigen binding site.

General Techniques

For further elaboration of general techniques useful in the practice of this invention, the practitioner can refer to standard textbooks and reviews in cell biology, tissue culture, and embryology. Included are *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach* (E.J. Robertson, ed., IRL Press Ltd. 1987); *Guide to Techniques in Mouse Development* (P.M. Wasserman et al., eds., Academic Press 1993); *Embryonic Stem Cell Differentiation in Vitro* (M.V. Wiles, Meth. Enzymol. 225:900, 1993); *Properties and uses of Embryonic Stem Cells: Prospects for Application to Human Biology and Gene Therapy* (P.D. Rathjen et al., Reprod. Fertil. Dev. 10:31, 1998). Differentiation of stem cells is reviewed in Robertson, Meth. Cell Biol. 75:173, 1997; and Pedersen, Reprod. Fertil. Dev. 10:31, 1998.

Methods in molecular genetics and genetic engineering are described generally in the current editions of *Molecular Cloning: A Laboratory Manual*, (Sambrook et al.); *Oligonucleotide Synthesis* (M.J. Gait, ed.); *Animal Cell Culture* (R.I. Freshney, ed.); *Gene Transfer Vectors for Mammalian Cells* (Miller & Calos, eds.); *Current Protocols in Molecular Biology* and *Short Protocols in Molecular Biology, 3rd Edition* (F.M. Ausubel et al., eds.); and *Recombinant DNA Methodology* (R. Wu ed., Academic Press). Reagents, cloning vectors, and kits for genetic manipulation referred to in this disclosure are available from commercial vendors such as BioRad, Stratagene, Invitrogen, and ClonTech.

General techniques in cell culture and media collection are outlined in *Large Scale Mammalian Cell Culture* (Hu et al., Curr. Opin. Biotechnol. 8:148, 1997); *Serum-free Media* (K. Kitano, Biotechnology 17:73, 1991); *Large Scale Mammalian Cell Culture* (Curr. Opin. Biotechnol. 2:375, 1991); and *Suspension Culture of Mammalian Cells* (Birch et al., Bioprocess Technol. 19:251, 1990).

Sources of Stem Cells

This invention can be practiced using stem cells of various types, which may include the following non-limiting examples.

U.S. Patent 5,851,832 reports multipotent neural stem cells obtained from brain tissue. U.S. Patent 5,766,948 reports producing neuroblasts from newborn cerebral hemispheres. U.S. Patent 5,654,183 and 5,849,553 report the use of mammalian neural crest stem cells. U.S. Patent 6,040,180 reports in vitro generation of differentiated neurons from cultures of mammalian multipotential CNS stem cells. WO 98/50526 and WO 99/01159 report generation and isolation of neuroepithelial stem cells, oligodendrocyte-astrocyte precursors, and lineage-restricted neuronal precursors. U.S. Patent 5,968,829 reports neural stem cells obtained from embryonic forebrain and cultured with a medium comprising glucose, transferrin, insulin, selenium, progesterone, and several other growth factors.

Primary liver cell cultures can be obtained from human biopsy or surgically excised tissue by perfusion with an appropriate combination of collagenase and hyaluronidase. Alternatively, EP 0 953 633 A1 reports

isolating liver cells by preparing minced human liver tissue, resuspending concentrated tissue cells in a growth medium and expanding the cells in culture. The growth medium comprises glucose, insulin, transferrin, T₃, FCS, and various tissue extracts that allow the hepatocytes to grow without malignant transformation. The cells in the liver are thought to contain specialized cells including liver parenchymal cells, Kupffer cells, sinusoidal endothelium, and bile duct epithelium, and also precursor cells (referred to as "hepatoblasts" or "oval cells") that have the capacity to differentiate into both mature hepatocytes or biliary epithelial cells (L.E. Rogler, Am. J. Pathol. 150:591, 1997; M. Alison, Current Opin. Cell Biol. 10:710, 1998; Lazaro et al., Cancer Res. 58:514, 1998).

U.S. Patent 5,192,553 reports methods for isolating human neonatal or fetal hematopoietic stem or progenitor cells. U.S. Patent 5,716,827 reports human hematopoietic cells that are Thy-1 positive progenitors, and appropriate growth media to regenerate them *in vitro*. U.S. Patent 5,635,387 reports a method and device for culturing human hematopoietic cells and their precursors. U.S. Patent 6,015,554 describes a method of reconstituting human lymphoid and dendritic cells.

U.S. Patent 5,486,359 reports homogeneous populations of human mesenchymal stem cells that can differentiate into cells of more than one connective tissue type, such as bone, cartilage, tendon, ligament, and dermis. They are obtained from bone marrow or periosteum. Also reported are culture conditions used to expand mesenchymal stem cells. WO 99/01145 reports human mesenchymal stem cells isolated from peripheral blood of individuals treated with growth factors such as G-CSF or GM-CSF. WO 00/53795 reports adipose-derived stem cells and lattices, substantially free of adipocytes and red cells. These cells reportedly can be expanded and cultured to produce hormones and conditioned culture media.

The invention can be practiced using stem cells of any vertebrate species. Included are stem cells from humans; as well as non-human primates, domestic animals, livestock, and other non-human mammals.

Amongst the stem cells suitable for use in this invention are primate pluripotent stem (pPS) cells derived from tissue formed after gestation, such as a blastocyst, or fetal or embryonic tissue taken any time during gestation. Non-limiting examples are primary cultures or established lines of embryonic stem cells.

Media and Feeder Cells

Media for isolating and propagating pPS cells can have any of several different formulas, as long as the cells obtained have the desired characteristics, and can be propagated further. Suitable sources are as follows: Dulbecco's modified Eagles medium (DMEM), Gibco # 11965-092; Knockout Dulbecco's modified Eagles medium (KO DMEM), Gibco # 10829-018; 200 mM L-glutamine, Gibco # 15039-027; non-essential amino acid solution, Gibco 11140-050; β -mercaptoethanol, Sigma # M7522; human recombinant basic fibroblast growth factor (bFGF), Gibco # 13256-029. Exemplary serum-containing ES medium is made with 80% DMEM (typically KO DMEM), 20% defined fetal bovine serum (FBS) not heat inactivated, 0.1 mM non-essential amino acids, 1 mM L-glutamine, and 0.1 mM β -mercaptoethanol. The medium is filtered and stored at 4°C for no longer than 2 weeks. Serum-free ES medium is made with 80% KO DMEM, 20% serum replacement, 0.1 mM non-essential amino acids, 1 mM L-glutamine, and 0.1 mM β -mercaptoethanol. An effective serum replacement is Gibco # 10828-028. The medium is filtered and stored at 4°C for no longer than 2 weeks. Just before use, human bFGF is added to a final concentration of 4 ng/mL (Bodnar et al., Geron Corp, International Patent Publication WO 99/20741).

Feeder cells (where used) are propagated in mEF medium, containing 90% DMEM (Gibco # 11965-092), 10% FBS (Hyclone # 30071-03), and 2 mM glutamine. mEFs are propagated in T150 flasks (Corning # 430825), splitting the cells 1:2 every other day with trypsin, keeping the cells subconfluent. To prepare the feeder cell layer, cells are irradiated at a dose to inhibit proliferation but permit synthesis of important factors that support hES cells

(~4000 rads gamma irradiation). Six-well culture plates (such as Falcon # 304) are coated by incubation at 37°C with 1 mL 0.5% gelatin per well overnight, and plated with 375,000 irradiated mEFs per well. Feeder cell layers are typically used 5 h to 4 days after plating. The medium is replaced with fresh hES medium just before seeding pPS cells.

- 5 Conditions for culturing other stem cells are known, and can be optimized appropriately according to the cell type. Media and culture techniques for particular cell types referred to in the previous section are provided in the references cited.

Embryonic Stem Cells

- 10 Embryonic stem cells can be isolated from blastocysts of members of the primate species (Thomson et al., Proc. Natl. Acad. Sci. USA 92:7844, 1995). Human embryonic stem (hES) cells can be prepared from human blastocyst cells using the techniques described by Thomson et al. (U.S. Patent 5,843,780; Science 282:1145, 1998; Curr. Top. Dev. Biol. 38:133 ff., 1998) and Reubinoff et al, Nature Biotech. 18:399,2000.

- 15 Briefly, human blastocysts are obtained from human in vivo preimplantation embryos. Alternatively, in vitro fertilized (IVF) embryos can be used, or one cell human embryos can be expanded to the blastocyst stage (Bongso et al., Hum Reprod 4: 706, 1989). Human embryos are cultured to the blastocyst stage in G1.2 and G2.2 medium (Gardner et al., Fertil. Steril. 69:84, 1998). Blastocysts that develop are selected for ES cell isolation. The zona pellucida is removed from blastocysts by brief exposure to pronase (Sigma). The inner cell masses are isolated by immunosurgery, in which blastocysts are exposed to a 1:50 dilution of rabbit anti-human spleen cell antiserum for 30 minutes, then washed for 5 minutes three times in DMEM, and exposed to a 1:5 dilution of Guinea pig complement (Gibco) for 3 minutes (see Solter et al., Proc. Natl. Acad. Sci. USA 72:5099, 1975). After two further washes in DMEM, lysed trophoctoderm cells are removed from the intact inner cell mass (ICM) by gentle pipetting, and the ICM plated on mEF feeder layers.

- 20 After 9 to 15 days, inner cell mass-derived outgrowths are dissociated into clumps either by exposure to calcium and magnesium-free phosphate-buffered saline (PBS) with 1 mM EDTA, by exposure to dispase or trypsin, or by mechanical dissociation with a micropipette; and then replated on mEF in fresh medium. Dissociated cells are replated on mEF feeder layers in fresh ES medium, and observed for colony formation. Colonies demonstrating undifferentiated morphology are individually selected by micropipette, mechanically dissociated into clumps, and replated. ES-like morphology is characterized as compact colonies with apparently high nucleus to cytoplasm ratio and prominent nucleoli. Resulting ES cells are then routinely split every 1-2 weeks by brief trypsinization, exposure to Dulbecco's PBS (without calcium or magnesium and with 2 mM EDTA), exposure to type IV collagenase (~200 U/mL; Gibco) or by selection of individual colonies by micropipette. Clump sizes of about 50 to 100 cells are optimal.

Embryonic Germ Cells

35 Human Embryonic Germ (hEG) cells can be prepared from primordial germ cells present in human fetal material taken about 8-11 weeks after the last menstrual period. Suitable preparation methods are described in Shambloott et al., Proc. Natl. Acad. Sci. USA 95:13726, 1998 and U.S. Patent 6,090,622.

- 40 Briefly, genital ridges are rinsed with isotonic buffer, then placed into 0.1mL 0.05% trypsin/0.53 mM sodium EDTA solution (BRL) and cut into <1 mm³ chunks. The tissue is then pipetted through a 100 µL tip to further disaggregate the cells. It is incubated at 37°C for ~5 min, then ~3.5 mL EG growth medium is added. EG growth medium is DMEM, 4500 mg/L D-glucose, 2200 mg/L mM sodium bicarbonate; 15% ES qualified fetal calf

serum (BRL); 2 mM glutamine (BRL); 1 mM sodium pyruvate (BRL); 1000-2000 U/mL human recombinant leukemia inhibitory factor (LIF, Genzyme); 1-2 ng/ml human recombinant basic fibroblast growth factor (bFGF, Genzyme); and 10 μ M forskolin (in 10% DMSO). In an alternative approach, EG cells are isolated using hyaluronidase/collagenase/DNase. Gonadal anlagen or genital ridges with mesenteries are dissected from fetal material, the genital ridges are rinsed in PBS, then placed in 0.1 mL HCD digestion solution (0.01 % hyaluronidase type V, 0.002% DNase I, 0.1% collagenase type IV, all from Sigma prepared in EG growth medium). Tissue is minced and incubated 1 h or overnight at 37°C, resuspended in 1-3 mL of EG growth medium, and plated onto a feeder layer.

Ninety-six well tissue culture plates are prepared with a sub-confluent layer of feeder cells cultured for 3 days in modified EG growth medium free of LIF, bFGF or forskolin, inactivated with 5000 rad γ -irradiation. Suitable feeders are STO cells (ATCC Accession No. CRL 1503). ~0.2 mL of primary germ cell (PGC) suspension is added to each of the wells. The first passage is conducted after 7-10 days in EG growth medium, transferring each well to one well of a 24-well culture dish previously prepared with irradiated STO mouse fibroblasts. The cells are cultured with daily replacement of medium until cell morphology consistent with EG cells are observed, typically after 7-30 days or 1-4 passages.

Propagation of pPS Cells in an Undifferentiated State

pPS cells can be propagated continuously in culture, using a combination of culture conditions that promote proliferation without promoting differentiation.

Traditionally, pPS cells are cultured on a layer of feeder cells, typically fibroblast type cells, often derived from embryonic or fetal tissue. The cell lines are plated to near confluence, usually irradiated to prevent proliferation, and then used to support pPS cell cultures.

In one illustration, pPS cells are first derived and supported on primary embryonic fibroblasts. Mouse embryonic fibroblasts (mEF) can be obtained from outbred CF1 mice (SASCO) or other suitable strains. The abdomen of a mouse at 13 days of pregnancy is swabbed with 70% ethanol, and the decidua is removed into phosphate buffered saline (PBS). Embryos are harvested; placenta, membranes, and soft tissues are removed; and the carcasses are washed twice in PBS. They are then transferred to fresh 10 cm bacterial dishes containing 2 mL trypsin/EDTA, and finely minced. After incubating 5 min at 37°C, the trypsin is inactivated with 5 mL DMEM containing 10% bovine serum (FBS), and the mixture is transferred to a 15 mL conical tube and dissociated. Debris is allowed to settle for 2 min, the supernatant is made up to a final volume of 10 mL, and plated onto a 10 cm tissue culture plate or T75 flask. The flask is incubated undisturbed for 24 h, after which the medium is replaced. When flasks are confluent (~2-3 d), they are split 1:2 into new flasks.

Scientists at Geron have discovered that hPS cells can be maintained in an undifferentiated state even without feeder cells. The environment for feeder-free cultures includes a suitable culture substrate, particularly an extracellular matrix, such as may be derived from basement membrane or that may form part of adhesion molecule receptor-ligand couplings. A suitable preparation available from Becton Dickinson under the name Matrigel®. Other extracellular matrix components and component mixtures are suitable as an alternative. Depending on the cell type being proliferated, this may include laminin, fibronectin, proteoglycan, entactin, heparan sulfate, and the like, alone or in various combinations. Laminins are major components of all basal laminae in vertebrates, which interact with integrin heterodimers such as $\alpha 6 \beta 1$ and $\alpha 6 \beta 4$ (specific for laminins) and other heterodimers (that cross-react with other matrices).

The pluripotent stem cells are plated onto the substrate in a suitable distribution and in the presence of a medium that promotes cell survival, propagation, and retention of the desirable characteristics. It has been found that plating densities of at least $\sim 15,000$ cells cm^{-2} (typically $90,000$ cm^{-2} to $170,000$ cm^{-2}) promote survival and limit differentiation. The passage of pPS cells in the absence of feeders benefits from preparing the pPS cells in small clusters. Typically, enzymatic digestion is halted before cells become completely dispersed (say, ~ 5 min with collagenase IV). Clumps of ~ 10 -2000 cells are then plated directly onto the substrate without further dispersal.

Alternatively, primate PS cells can be passaged between feeder-free cultures as a finer cell suspension, providing that an appropriate enzyme and medium are chosen, and the plating density is sufficiently high. By way of illustration, confluent human embryonic stem cells cultured in the absence of feeders are removed from the plates by incubating with a solution of 0.05% (wt/vol) trypsin (Gibco) and 0.053 mM EDTA for 5-15 min at 37°C . With the use of pipette, the remaining cells in the plate are removed and the cells are triturated with the pipette until the cells are dispersed into a suspension comprising single cells and some small clusters. The cells are then plated at densities of $50,000$ - $200,000$ cells/ cm^2 to promote survival and limit differentiation. The phenotype of ES cells passaged by this technique is similar to what is observed when cells are harvested as clusters by collagen digestion. As another option, the cells can be harvested without enzymes before the plate reaches confluence. The cells are incubated ~ 5 min in a solution of 0.5 mM EDTA alone in PBS, washed from the culture vessel, and then plated into a new culture without further dispersal.

pPS cells plated in the absence of fresh feeder cells benefit from being cultured in a nutrient medium. The medium will generally contain the usual components to enhance cell survival, including isotonic buffer, essential minerals, and either serum or a serum replacement of some kind. Conditioned medium can be prepared by culturing irradiated primary mouse embryonic fibroblasts (or another suitable cell preparation) at a density of ~ 5 - 6×10^4 cm^{-2} in a serum free medium such as KO DMEM supplemented with 20% serum replacement and 4 ng/mL basic fibroblast growth factor (bFGF). The culture supernatant is harvested after ~ 1 day at 37°C .

As an alternative to primary mouse fibroblast cultures, conditioned medium can be prepared from an embryonic fibroblast cell line tested for its ability to condition medium appropriately. Such lines can optionally be transfected with telomerase reverse transcriptase to increase their replicative capacity. Another possible source is differentiated pPS cells with the morphological features of fibroblasts. pPS cells are suspension cultured as aggregates in differentiation medium using non-adherent cell culture plates ($\sim 2 \times 10^6$ cells/ 9.6 cm^2). After 2 days the aggregates are transferred into gelatin-coated plates, and fibroblast-like cells appear in clusters of 100-1000 cells in the mixed population after ~ 11 days. After brief collagenase treatment, the fibroblast-like cells can be collected under a microscope, passaged in mEF medium, and tested for their ability to condition ES medium.

Medium that has been conditioned for 1-2 days is typically used to support pPS cell culture for 1-2 days, and then exchanged. If desired, conditioned medium can be supplemented before use with additional growth factors that benefit pPS cell culture. For hES, a growth factor like bFGF or FGF-4 can be used. For hEG, culture medium may be supplemented with a growth factor like bFGF, an inducer of gp130, such as LIF or Oncostatin-M, and perhaps a factor that elevates cyclic AMP levels, such as forskolin.

Characteristics of Undifferentiated pPS Cells

In the two dimensions of a standard microscopic image, hES cells have high nuclear/cytoplasmic ratios in the plane of the image, prominent nucleoli, and compact colony formation with poorly discernable cell junctions. Cell lines can be karyotyped using a standard G-banding technique (available at many clinical diagnostics labs that

provides routine karyotyping services, such as the Cytogenetics Lab at Oakland CA) and compared to published human karyotypes.

hES and hEG cells can also be characterized on the basis of expressed cell markers. In general, the tissue-specific markers discussed in this disclosure can be detected using a suitable immunological technique — such as flow cytometry for membrane-bound markers, immunohistochemistry for intracellular markers, and enzyme-linked immunoassay, for markers secreted into the medium. The expression of protein markers can also be detected at the mRNA level by reverse transcriptase-PCR using marker-specific primers. See U.S. Patent 5,843,780 for further details.

Stage-specific embryonic antigens (SSEA) are characteristic of certain embryonic cell types. Antibodies for SSEA markers are available from the Developmental Studies Hybridoma Bank (Bethesda MD). Other useful markers are detectable using antibodies designated Tra-1-60 and Tra-1-81 (Andrews et al., *Cell Lines from Human Germ Cell Tumors*, in E.J. Robertson, 1987, *supra*). hES cells are typically SSEA-1 negative and SSEA-4 positive. hEG cells are typically SSEA-1 positive. Differentiation of pPS cells in vitro results in the loss of SSEA-4, Tra-1-60, and Tra-1-81 expression and increased expression of SSEA-1. pPS cells can also be characterized by the presence of alkaline phosphatase activity, which can be detected by fixing the cells with 4% paraformaldehyde, and then developing with Vector Red as a substrate, as described by the manufacturer (Vector Laboratories, Burlingame CA).

Embryonic stem cells are also typically telomerase positive and OCT-4 positive. Telomerase activity can be determined using TRAP activity assay (Kim et al., *Science* 266:2011, 1997), using a commercially available kit (TRAPeze® XK Telomerase Detection Kit, Cat. s7707; Intergen Co., Purchase NY; or TeloTAGGG™ Telomerase PCR ELISApplus, Cat. 2,013,89; Roche Diagnostics, Indianapolis). hTERT expression can also be evaluated at the mRNA level by RT-PCR. The LightCycler TeloTAGGG™ hTERT quantification kit (Cat. 3,012,344; Roche Diagnostics) is available commercially for research purposes.

Differentiating pPS Cells

Differentiation of the pPS can be initiated by first forming embryoid bodies. General principles in culturing embryoid bodies are reported in O'Shea, *Anat. Rec. (New Anat. 257:323, 1999)*. pPS cells are cultured in a manner that permits aggregates to form, for which many options are available: for example, by overgrowth of a donor pPS cell culture, or by culturing pPS cells in culture vessels having a substrate with low adhesion properties which allows EB formation. Embryoid bodies can also be made in suspension culture. pPS cells are harvested by brief collagenase digestion, dissociated into clusters, and plated in non-adherent cell culture plates. The aggregates are fed every few days, and then harvested after a suitable period, typically 4-8 days. The cells can then be cultured in a medium and/or on a substrate that promotes enrichment of cells of a particular lineage. The substrate can comprise matrix components such as Matrigel® (Becton Dickenson), laminin, collagen, gelatin, or matrix produced by first culturing a matrix-producing cell line (such as a fibroblast or endothelial cell line), and then lysing and washing in such a way that the matrix remains attached to the surface of the vessel. Embryoid bodies comprise a heterogeneous cell population, potentially having an endoderm exterior, and a mesoderm and ectoderm interior.

Scientists at Geron Corporation have discovered that pPS cells can be differentiated into committed precursor cells or terminally differentiated cells without forming embryoid bodies or aggregates as an intermediate step. Briefly, a suspension of undifferentiated pPS cells is prepared, and then plated onto a solid surface that promotes differentiation. Suitable substrates include glass or plastic surfaces that are adherent. For example,

glass coverslips can be coated with a polycationic substance, such as a polyamines like poly-lysine, poly-ornithine, or other homogeneous or mixed polypeptides or other polymers with a predominant positive charge. The cells are then cultured in a suitable nutrient medium that is adapted to promote differentiation towards the desired cell lineage.

In some circumstances, differentiation is further promoted by withdrawing serum or serum replacement from the culture medium. This can be achieved by substituting a medium devoid of serum and serum replacement, for example, at the time of replating. In certain embodiments of the invention, differentiation is promoted by withdrawing one or more medium component(s) that promote(s) growth of undifferentiated cells, or act(s) as an inhibitor of differentiation. Examples of such components include certain growth factors, mitogens, leukocyte inhibitory factor (LIF), and basic fibroblast growth factor (bFGF). Differentiation may also be promoted by adding a medium component that promotes differentiation towards the desired cell lineage, or inhibits the growth of cells with undesired characteristics. For example, to generate cells committed to neural or glial lineages, the medium can include any of the following factors or medium constituents in an effective combination: Brain derived neurotrophic factor (BDNF), neurotrophin-3 (NT-3), NT-4, epidermal growth factor (EGF), ciliary neurotrophic factor (CNTF), nerve growth factor (NGF), retinoic acid (RA), sonic hedgehog, FGF-8, ascorbic acid, forskolin, fetal bovine serum (FBS), and bone morphogenic proteins (BMPs).

General principals for obtaining tissue cells from pluripotent stem cells are reviewed in Pedersen (Reprod. Fertil. Dev. 6:543, 1994), and U.S. Patent 6,090,622. Other publications of interest include the following: For neural progenitors, neural restrictive cells and glial cell precursors, see Bain et al., Biochem. Biophys. Res. Commun. 200:1252, 1994; Trojanowski et al., Exp. Neurol. 144:92, 1997; Wojcik et al., Proc. Natl. Acad. Sci. USA 90:1305-130; and U.S. Patents 5,851,832, 5,928,947, 5,766,948, and 5,849,553. For cardiac muscle and cardiomyocytes see Chen et al., Dev. Dynamics 197:217, 1993 and Wobus et al., Differentiation 48:173, 1991. For hematopoietic progenitors, see Burkert et al., New Biol. 3:698, 1991 and Biesecker et al., Exp. Hematol. 21:774, 1993. U.S. Patent 5,773,255 relates to glucose-responsive insulin secreting pancreatic beta cell lines. U.S. Patent 5,789,246 relates to hepatocyte precursor cells. Other progenitors of interest include but are not limited to chondrocytes, osteoblasts, retinal pigment epithelial cells, fibroblasts, skin cells such as keratinocytes, dendritic cells, hair follicle cells, renal duct epithelial cells, smooth and skeletal muscle cells, testicular progenitors, and vascular endothelial cells.

Scientists at Geron Corporation have discovered that culturing pPS cells or embryoid body cells in the presence of ligands that bind growth factor receptors promotes enrichment for neural precursor cells. The growth environment may contain a neural cell supportive extracellular matrix, such as fibronectin. Suitable growth factors include but are not limited to EGF, bFGF, PDGF, IGF-1, and antibodies to receptors for these ligands. The cultured cells may then be optionally separated on the basis of whether they express a marker such as A2B5. Under the appropriate circumstances, populations of cells enriched for expression of the A2B5 marker may have the capacity to generate both neuronal cells (including mature neurons), and glial cells (including astrocytes and oligodendrocytes). Optionally, the cell populations are further differentiated, for example, by culturing in a medium containing an activator of cAMP.

Scientists at Geron Corporation have also discovered that culturing pPS cells or embryoid body cells in the presence of a hepatocyte differentiation agent promotes enrichment for hepatocyte-like cells. The growth environment may contain a hepatocyte supportive extracellular matrix, such as collagen or Matrigel®. Suitable differentiation agents include various isomers of butyrate and their analogs, exemplified by n-butyrate. The cultured cells are optionally cultured simultaneously or sequentially with a hepatocyte maturation factor, such as an

organic solvent like dimethyl sulfoxide (DMSO); a maturation cofactor such as retinoic acid; or a cytokine or hormone such as a glucocorticoid, epidermal growth factor (EGF), insulin, TGF- α , TGF- β , fibroblast growth factor (FGF), heparin, hepatocyte growth factor (HGF), IL-1, IL-6, IGF-I, IGF-II, and HBGF-1.

5 *Characteristics of differentiated cells*

Cells can be characterized according to a number of phenotypic criteria. The criteria include but are not limited to characterization of morphological features, detection or quantitation of expressed cell markers and enzymatic activity, and determination of the functional properties of the cells in vivo.

10 Markers of interest for neural cells include β -tubulin III or neurofilament, characteristic of neurons; glial fibrillary acidic protein (GFAP), present in astrocytes; galactocerebroside (GalC) or myelin basic protein (MBP); characteristic of oligodendrocytes; OCT-4, characteristic of undifferentiated hES cells; nestin, characteristic of neural precursors and other cells. A2B5 and NCAM are characteristic of glial progenitors and neural progenitors, respectively. Cells can also be tested for secretion of characteristic biologically active substances. For example, GABA-secreting neurons can be identified by production of glutamic acid decarboxylase or GABA. Dopaminergic
15 neurons can be identified by production of dopa decarboxylase, dopamine, or tyrosine hydroxylase.

Markers of interest for liver cells include α -fetoprotein (liver progenitors); albumin, α_1 -antitrypsin, glucose-6-phosphatase, cytochrome p450 activity, transferrin, asialoglycoprotein receptor, and glycogen storage (hepatocytes); CK7, CK19, and γ -glutamyl transferase (bile epithelium). It has been reported that hepatocyte differentiation requires the transcription factor HNF-4 α (Li et al., Genes Dev. 14:464, 2000). Markers independent
20 of HNF-4 α expression include α_1 -antitrypsin, α -fetoprotein, apoE, glucokinase, insulin growth factors 1 and 2, IGF-1 receptor, insulin receptor, and leptin. Markers dependent on HNF-4 α expression include albumin, apoA1, apoAII, apoB, apoCIII, apoCII, aldolase B, phenylalanine hydroxylase, L-type fatty acid binding protein, transferrin, retinol binding protein, and erythropoietin (EPO).

Cell types in mixed cell populations derived from pPS cells can be recognized by characteristic
25 morphology and the markers they express. For skeletal muscle: myoD, myogenin, and myf-5. For endothelial cells: PECAM (platelet endothelial cell adhesion molecule), Flk-1, tie-1, tie-2, vascular endothelial (VE) cadherin, MECA-32, and MEC-14.7. For smooth muscle cells: specific myosin heavy chain. For cardiomyocytes: GATA-4, Nkx2.5, cardiac troponin I, α -myosin heavy chain, and ANF. For pancreatic cells, pdx and insulin secretion. For hematopoietic cells and their progenitors: GATA-1, CD34, AC133, β -major globulin, and β -major globulin like gene
30 β H1.

Certain tissue-specific markers listed in this disclosure or known in the art can be detected by immunological techniques — such as flow immunocytochemistry for cell-surface markers, immunohistochemistry (for example, of fixed cells or tissue sections) for intracellular or cell-surface markers, Western blot analysis of cellular extracts, and enzyme-linked immunoassay, for cellular extracts or products secreted into the medium. The
35 expression of tissue-specific gene products can also be detected at the mRNA level by Northern blot analysis, dot-blot hybridization analysis, or by reverse transcriptase initiated polymerase chain reaction (RT-PCR) using sequence-specific primers in standard amplification methods. Sequence data for the particular markers listed in this disclosure can be obtained from public databases such as GenBank (URL www.ncbi.nlm.nih.gov/80/entrez).

Preparing Cell Populations Essentially Free of Undifferentiated Cells

In accordance with this invention, populations of differentiated cells are depleted of relatively undifferentiated cells by expressing a gene that is lethal to cells or renders them susceptible to a lethal effect of an external agent, under control of a transcriptional control element that causes the gene to be preferentially expressed in the undifferentiated cells.

To accomplish this, the cells are genetically altered either before or after the process used to differentiate the cells into the desired lineage for therapy, in a way that puts an effector gene suitable for negative selection of undifferentiated cells, under control of a transcriptional control element with the desired properties.

Transcriptional Control Elements for Driving Negative Selection

The control element is selected with a view to the protein expression patterns of the undifferentiated and differentiated cells in the population.

Genes with desirable expression patterns can be identified by comparing expression at the transcription, translation, or functional level in two different cell populations — one relatively enriched for differentiated cells, the other relatively enriched for undifferentiated cells. Suitable methods of comparison include subtractive hybridization of cDNA libraries, and microarray analysis of mRNA levels. Once a transcript is identified with an appropriate expression pattern, the promoter or enhancer of the corresponding gene can be used for construction of the negative selection vector.

A suitable microarray analysis is conducted using a Genetic Microsystems array generator, and an Axon GenePix™ Scanner. Microarrays are prepared by amplifying cDNA fragments in a 96 or 384 well format, and then spotted directly onto glass slides. To compare mRNA preparations from two cell populations, one preparation is converted into Cy3-labeled cDNA, while the other is converted into Cy5-labeled cDNA. The two cDNA preparations are hybridized simultaneously to the microarray slide, and then washed to eliminate non-specific binding. Any given spot on the array will bind each of the cDNA products in proportion to abundance of the transcript in the two original mRNA preparations. The slide is then scanned at wavelengths appropriate for each of the labels, and the relative abundance of mRNA is determined. Preferably, the level of expression of the effector gene will be at least 5-fold or even 25-fold higher in the undifferentiated cells relative to the differentiated cells.

For the depletion of pluripotent embryonic cells, an exemplary control element is the promoter for telomerase reverse transcriptase (TERT). Sequence of the human TERT gene (including upstream promoter sequence) is provided below. The reader is also referred to U.K. Patent GB 2321642 B (Cech et al., Geron Corporation and U. Colorado), International Patent Publications WO 00/46355 (Morin et al., Geron Corporation), WO 99/33998 (Hagen et al., Bayer Aktiengesellschaft), and Horikawa, I., et al. (Cancer Res., 59:826, 1999). Sequence of the mouse TERT gene is provided in WO 99/27113 (Morin et al., Geron Corporation). A lambda phage clone designated λGΦ5, containing ~13,500 bases upstream from the hTERT encoding sequence is available from the ATCC under Accession No. 98505. Example 9 illustrates the testing and use of TERT promoter sequences (SEQ. ID NO:1) in vector expression systems.

Another exemplary control element is a promoter sequence for Octamer binding transcription factor 4 (OCT-4), a member of the POU family of transcription factors. OCT-4 transcription is activated between the 4 and 80 cell stage in the developing embryo, and it is highly expressed in the expanding blastocyst and then in the pluripotent cells of the egg cylinder. Transcription is down-regulated as the primitive ectoderm differentiates to form mesoderm, and by 8.5 days post coitum is restricted to migrating primordial germ cells. High-level OCT-4 gene expression is also observed in pluripotent embryo carcinoma and embryonic stem cell lines, and is down-

regulated when these cells are induced to differentiate. Pig, mouse, and human OCT-4 promoter sequences are provided in International Patent Publication WO 9919469 (Biotransplant Inc.).

Other suitable control elements can be obtained from genes causing expression of markers characteristic of undifferentiated cells in the population but not of the differentiated cells. For example, SSEA-3, SSEA-4, Tra-1-60 and Tra-1-81 are characteristic of various types of undifferentiated pluripotent embryonic stem cells. The enzyme responsible for synthesis of SSEA-4 may have transcriptional control elements with the desirable expression specificity. A more recent example is the promoter for *Rex1* protein, a retanoic acid regulated zinc finger protein that is expressed in preimplantation embryos. The mouse *Rex1* promoter has been shown to act as an effective transcription marker for undifferentiated embryonic stem cells (Eiges et al., Current Biol. 11:514, 2001).

Suitability of particular elements can be estimated by analysis of gene transcript expression, for example, by microarray analysis. Reporter constructs can then be tested in differentiated and undifferentiated cells for the appropriate specificity, using a promoter or enhancer sequence from the identified cell-specific gene to control transcription of a reporter gene, such as green fluorescence protein, secreted alkaline phosphatase, β -glucuronidase, or β -galactosidase. Use of reporter constructs to test promoter specificity is illustrated below in Example 9.

Effector Genes for Achieving Negative Selection

A transcriptional regulatory element with appropriate specificity is operatively linked to an encoding region for a product that will provide elimination of cells in which it is expressed — either directly, or by rendering the cell susceptible to an otherwise innocuous external agent.

Of particular interest are genes that cause presentation of a foreign antigen on the cell membrane. The presented substance may be an alloantigen, a xenoantigen, or an antigen from a non-mammalian species for which specific antibody is readily available. Expression of the gene leads to presentation of the antigen on undifferentiated cells, which then can be used to effect depletion by a suitable immunological separation.

In one embodiment, the effector sequence encodes a membrane protein that contains the epitope recognized by the specific antibody. The membrane protein may be a protein expressed in the same species on other types of cells, but more typically is obtained from another species, or is an artificial sequence. In this case, the antigen will be foreign to the species from which the stem cells are derived, and antibodies made in the same species will not cross-react with other antigens on the cell. Included are xenoantigens, alloantigens, and artificial antigens (made by constructing an immunogenic peptide sequence not encoded in the human genome).

In another embodiment, the target antigen is not a protein, but a carbohydrate or lipid component. In this case, the effector sequence will encode an enzyme involved in antigen synthesis. Of particular interest are glycosyl transferases of mammalian or non-mammalian origin that synthesize carbohydrate differentiation antigen, alloantigens, xenoantigen, or novel determinants detectable by antibody. Examples include the following:

- The marker SSEA-1, for which the effector sequence encodes the corresponding fucosyltransferase.
- The Gal α (1,3)Gal linkage present on endothelial tissue of most mammals except for humans and old-world monkeys, formed by an α (1,3)galactosyltransferase (α 1,3GT). The encoding sequence for sheep and marmoset α 1,3GT are provided below (see also Henion et al., Glycobiology 4:193, 1994). In the processing of human cells, it is possible to derive a human equivalent α 1,3GT, by correcting the silent human α 1,3GT pseudogene (Joziasse et al., J. Biol. Chem. 266:6991, 1991) using a consensus of the lower primate α 1,3GT sequences.

- The ABO histo blood group antigens present on most human cells. The encoding sequence is the corresponding ABO transferases, for which the encoding sequences are provided below.

Cells displaying the target antigen are separated using a ligand such as an antibody or lectin specific for the target. General techniques used in producing antibodies and using them in immunoisolation are described in *Handbook of Experimental Immunology* (Weir & Blackwell, eds.); *Current Protocols in Immunology* (Coligan et al., eds.); and *Methods of Immunological Analysis* (Masseyeff et al., eds., Weinheim: VCH Verlags GmbH). Polyclonal antibodies can be prepared by injecting a vertebrate with the isolated membrane protein or an immunogenic fragment in a suitable adjuvant. Any unwanted cross-reactivity (such as reactivity against proteins expressed by the differentiated cells) can be removed by adsorbing with cross-reacting antigens attached to a solid phase, and collecting the unbound fraction. Production of monoclonal antibodies is described in such standard references as Harrow & Lane (1988), U.S. Patents 4,491,632, 4,472,500 and 4,444,887, and *Methods in Enzymology* 73B:3 (1981). Briefly, a mammal is immunized, their splenocytes are immortalized, and clones are selected that produce antibody of the desired specificity. Other methods of obtaining specific antibody molecules involve contacting a library of immunocompetent cells or viral particles (such as a phage display library) with the target antigen, and growing out positively selected clones (Marks et al., New Eng. J. Med. 335:730, 1996; WO 94/13804; WO 92/01047; WO 90/02809; and McGuinness et al., Nature Biotechnol. 14:1449, 1996).

Where the antibody target is a carbohydrate blood group alloantigen or xenoantigen, there are often "naturally occurring" antibodies already present in the circulation of mammals lacking the antigen without them having been deliberately immunized. It is thought that naturally occurring antibodies arise from cross-reactive carbohydrates present in the diet. Humans have naturally occurring antibodies to the human ABO blood group, and to the Gal α (1,3)Gal xenoantigen. The "IB4" lectin from *Bandeiraea (Griffonia) simplicifolia* (Sigma Cat. L 3019) is specific for α -D-galactosyl residues and binds both the Gal α (1,3)Gal epitope, and B blood group substance. Lectins of appropriate specificity can substitute for antibodies in some affinity separation procedures.

Where the antibody target is a carbohydrate antigen shared between species, it may be difficult to obtain a high affinity antibody from another mammal — in which case, it may be preferable to obtain antibody from an avian species or a phage display library. This invention also contemplates artificial carbohydrate antigens, built by using an effector glycosyltransferase that is not of mammalian origin, but capable of modifying mammalian complex carbohydrates. Novel carbohydrate antigens may be particularly immunogenic for raising antibody useful in depleting the undifferentiated cells.

In order to effect separation, the cell population is cultured under conditions that allow the transcriptional control element to drive expression of the effector sequence in undifferentiated cells. The cell population is then subjected to immunoseparation by combining with an antibody or lectin, and then recovering cells that are not bound. The undifferentiated cells will be in the bound fraction, and will be depleted from the population recovered. One separation technique is immunoaffinity binding, in which the antibody is attached to a solid support, and the cells are passed over the adsorbent. Differentiated cells will be in the nonadherent fraction. Another separation technique is fluorescence-activated cell sorting, in which the cells are contacted with a fluorescently labeled primary or secondary antibody. The differentiated cells will be in the non-fluorescent cell sort. Another separation technique is complement-mediated lysis. The cells are combined with a complete (complement-fixing) antibody, and simultaneously or sequentially with complement obtained from fresh serum or a commercial source. The complement will lyse undifferentiated cells, depleting them from the population.

As an alternative to ligand targets, the effector gene can encode a peptide toxin — such as ricin, abrin, diphtheria, gelonin, *Pseudomonas* exotoxin A, and so on. Also suitable are genes that induce or mediate

apoptosis — such as the ICE-family of cysteine proteases, the Bcl-2 family of proteins, Bax, bclXs and caspases. Koga et al. (Hu. Gene Ther. 11:1397, 2000) propose a telomerase-specific gene therapy using the hTERT gene promoter linked to the apoptosis gene Caspase-8 (FLICE). Gu et al. (Cancer Res. 60:5359, 2000) reported a binary adenoviral system that induced Bax expression via the hTERT promoter.

Other suitable effectors encode polypeptides having activity that is not itself toxic to a cell, but renders the cell sensitive to an otherwise nontoxic compound — either by metabolically altering the cell, or by changing a non-toxic prodrug into a lethal drug. Exemplary is thymidine kinase (*tk*), such as may be derived from a herpes simplex virus, which converts the anti-herpetic agent ganciclovir (GCV) to a toxic product that interferes with DNA replication. International Patent Publications WO 98/14593 and WO 00/46355 (Geron Corporation) describe constructs comprising HSV *tk* under control of hTERT promoter sequences. When the transducing agent is a viral vector, the effector can be a viral gene required for replication of the virus. Essential genes for replication of adenovirus include the E4, E1a, E1b, and E2 regions. See International Patent Publication WO 00/46355 (Morin et al., Geron Corporation) for a description of lytic vectors that replicate in cells expressing TERT.

Another possible effector sequence encodes a nucleic acid for antisense, a ribozyme or RNA interference (RNAi) targeting transcripts essential for cell viability. In one illustration, the hTERT promoter drives RNAi that inactivates hypoxanthine-guanine phosphoribosyl transferase (HGPRT) or *tk*. Residual undifferentiated cells can be removed from the population by incubation with medium is supplemented with hypoxanthine, aminopterin, and thymidine (HAT medium).

The vector constructs for use in this invention can also contain a positive selection marker, such as an antibiotic resistance gene, that is also under control of the specific promoter. Exemplary is a vector having the configuration hTERT promoter — *effector gene* — IRES or 2A sequence — *neo*. This is designed so that both the effector and drug resistance gene are expressed under control of the promoter that drives preferential expression in undifferentiated cells.

Selection Techniques to Eliminate Undifferentiated Cells

To deplete differentiated cell populations of undifferentiated cells, the effector gene is selectively expressed in the undifferentiated cells.

This can be accomplished in several ways. In one embodiment, the population is genetically altered using a vector in which a transcriptional control element of the appropriate specificity is operatively linked to the effector gene. The genetic alteration may be transient (for example, using an adenovirus vector), meaning that the level of expression diminishes as the cells divide. This is suitable for generating differentiated cell populations that will be free of heterologous genes at the time of therapy. The genetic alteration may also be permanent (for example, using a retroviral vector), meaning that the alteration is inheritable by progeny of the initially altered cell. This is suitable for generating differentiated cell populations that will have an ongoing corrective function as they proliferate in vitro or in vivo, to eliminate any undifferentiated or dedifferentiated cells that arise in the population.

Any suitable expression vector can be used. Suitable viral vector systems for producing stem cells altered according to this invention can be prepared using commercially available virus components. Viral vectors comprising effector genes are generally described in the publications referenced in the last section. Alternatively, vector plasmids can be introduced into cells by electroporation, or using lipid/DNA complexes, such as those described in U.S. Patent Nos. 5,578,475; 5,627,175; 5,705,308; 5,744,335; 5,976,567; 6,020,202; and 6,051,429. Exemplary is the formulation Lipofectamine 2000™, available from Gibco/Life Technologies. Another exemplary

reagent is FuGENE™ 6 Transfection Reagent, a blend of lipids in non-liposomal form and other compounds in 80% ethanol, obtainable from Roche Diagnostics Corporation.

In another embodiment, the effector gene is placed under control of an endogenous transcriptional control element, such as the hTERT or OCT-4 promoter. This can be effected, for example, by homologous recombination, using a vector comprising the effector encoding sequence, flanked on one side by the transcriptional control element and other upstream genomic sequence, and flanked on the other side by downstream genomic sequence for the targeted gene. U.S. Patent Nos. 5,464,764 and 5,631,153 describe a double-selection strategy, in which two sequences homologous to the gene target flank a positive selection marker, and a negative selection marker is attached to the 3' terminal of the second flanking region. U.S. Patent 5,789,215 reports the use of homologous recombination targeting vectors for modifying the cell genome of mouse embryonic stem cells. Other information of interest for homologous recombination targeting can be found in U.S. Patents 5,589,369, 5,776,774, and 5,789,215.

If the effector gene directly causes cell lysis or apoptosis, then the population will be depleted of undifferentiated cells upon culturing the cells under conditions where the control element is expected to cause transcription of the gene. However, if the effector gene is not directly lethal, but renders the cell susceptible to the lethal effects of an external agent, then depletion will be postponed until the external agent is provided. For example, where the gene is a prodrug converting enzyme, then depletion is effected upon placing the cells in an environment containing the prodrug. Where the gene creates an antibody target, then depletion is effected upon placing the cells in an environment containing specific antibody, plus complement. The environment can be a culture vessel, in which case the agent can just be added to the culture medium at the requisite concentration. Alternatively or in addition, depletion can be performed in vivo, by administering the cell population to a subject, and simultaneously or sequentially administering the agent, if not already present. Where the gene is a glycosyltransferase that creates a xenoantigen or an alloantigen, then placing the cells in a subject will render the cells liable to lysis in situ by naturally occurring antibody and complement.

Cell populations in which the majority of cells are differentiated can be genetically modified according to these procedures to deplete undifferentiated cells. Alternatively, a precursor population of relatively undifferentiated cells can be genetically modified according to these procedures, and then differentiated. In this situation, it is more typical to use an effector gene that does not kill the cells immediately upon expression, but renders the cells susceptible to the lethal effect of some external agent. In one illustration, undifferentiated pPS cells grown in culture are transduced with an adenovirus vector in which the herpes thymidine kinase gene is under control of the hTERT promoter. The cells are optionally selected for positive transduction, either by incorporating a selectable marker in the construct, or by measuring expression of the transduced gene, and proliferated in culture. When differentiated cells are desired, the population is taken through a differentiation procedure (for example, to make hepatocyte or neuron precursors, as described earlier). They are then cultured under conditions that permit expression of the *tk* gene in the presence of ganciclovir.

Cell populations may be obtained using these techniques that are "depleted" of undifferentiated cells, which indicates any significant reduction in the proportion of undifferentiated cells present. After the procedure is effected, the proportion of undifferentiated cells may be decreased by 50% or even 90%. Depending on the control element and effector chosen, it may be possible to achieve differentiated cell populations that are "essentially free" of undifferentiated cells. This means that the population as a whole contains less than 1% of cells with the undifferentiated phenotype. Populations containing less than 0.2%, 0.05%, or 0.01% undifferentiated cells are increasingly more preferred. For pPS cells, the presence of undifferentiated cells can be determined by

counting cells expressing SSEA-4 by FACS analysis, or by counting cells expressing TERT or OCT-4 by fluorescence in-situ hybridization.

Use of Differentiated Cells

Cells prepared according to this invention can be used for a variety of commercially important research, diagnostic, and therapeutic purposes.

Because the cell populations of this invention are depleted of undifferentiated cells, they can be used to prepare antibodies and cDNA libraries that are specific for the differentiated phenotype. General techniques used in raising, purifying and modifying antibodies are outlined in the references provided above. General techniques involved in preparation of mRNA and cDNA libraries are described in *RNA Methodologies: A Laboratory Guide for Isolation and Characterization* (R.E. Farrell, Academic Press, 1998); *cDNA Library Protocols* (Cowell & Austin, eds., Humana Press); and *Functional Genomics* (Hunt & Livesey, eds., 2000).

Relatively homogeneous cell populations are particularly suited for use in drug screening and therapeutic applications.

Drug Screening

Differentiated pPS cells of this invention can be used to screen for factors (such as solvents, small molecule drugs, peptides, polynucleotides, and the like) or environmental conditions (such as culture conditions or manipulation) that affect the characteristics of differentiated cells.

In some applications, differentiated cells are used to screen factors that promote maturation, or promote proliferation and maintenance of such cells in long-term culture. For example, candidate maturation factors or growth factors are tested by adding them to pPS cells in different wells, and then determining any phenotypic change that results, according to desirable criteria for further culture and use of the cells.

Particular screening applications of this invention relate to the testing of pharmaceutical compounds in drug research. The reader is referred generally to the standard textbook "In vitro Methods in Pharmaceutical Research", Academic Press, 1997, and U.S. Patent 5,030,015). Assessment of the activity of candidate pharmaceutical compounds generally involves combining the differentiated cells of this invention with the candidate compound, determining any change in the morphology, marker phenotype, or metabolic activity of the cells that is attributable to the compound (compared with untreated cells or cells treated with an inert compound), and then correlating the effect of the compound with the observed change.

The screening may be done, for example, either because the compound is designed to have a pharmacological effect on certain cell types, or because a compound designed to have effects elsewhere may have unintended side effects. Two or more drugs can be tested in combination (by combining with the cells either simultaneously or sequentially), to detect possible drug-drug interaction effects. In some applications, compounds are screened initially for potential toxicity (Castell et al., pp. 375-410 in "In vitro Methods in Pharmaceutical Research," Academic Press, 1997). Cytotoxicity can be determined in the first instance by the effect on cell viability, survival, morphology, and expression or release of certain markers, receptors or enzymes. Effects of a drug on chromosomal DNA can be determined by measuring DNA synthesis or repair. [³H]-thymidine or BrdU incorporation, especially at unscheduled times in the cell cycle, or above the level required for cell replication, is consistent with a drug effect. Unwanted effects can also include unusual rates of sister chromatid exchange, determined by metaphase spread. The reader is referred to A. Vickers (PP 375-410 in "In vitro Methods in Pharmaceutical Research," Academic Press, 1997) for further elaboration.

Therapeutic Use

Differentiated cells of this invention can also be used for tissue reconstitution or regeneration in a human patient in need thereof. The cells are administered in a manner that permits them to graft to the intended tissue site and reconstitute or regenerate the functionally deficient area.

In one example, neural stem cells are transplanted directly into parenchymal or intrathecal sites of the central nervous system, according to the disease being treated. Grafts are done using single cell suspension or small aggregates at a density of 25,000-500,000 cells per μL (U.S. Patent 5,968,829). The efficacy of neural cell transplants can be assessed in a rat model for acutely injured spinal cord as described by McDonald et al. (Nat. Med. 5:1410, 1999). A successful transplant will show transplant-derived cells present in the lesion 2-5 weeks later, differentiated into astrocytes, oligodendrocytes, and/or neurons, and migrating along the cord from the lesioned end, and an improvement in gate, coordination, and weight-bearing.

Certain neural progenitor cells embodied in this invention are designed for treatment of acute or chronic damage to the nervous system. For example, excitotoxicity has been implicated in a variety of conditions including epilepsy, stroke, ischemia, Huntington's disease, Parkinson's disease and Alzheimer's disease. Certain differentiated cells of this invention may also be appropriate for treating dysmyelinating disorders, such as Pelizaeus-Merzbacher disease, multiple sclerosis, leukodystrophies, neuritis and neuropathies. Appropriate for these purposes are cell cultures enriched in oligodendrocytes or oligodendrocyte precursors to promote remyelination.

Hepatocytes and hepatocyte precursors prepared according to this invention can be assessed in animal models for ability to repair liver damage. One such example is damage caused by intraperitoneal injection of D-galactosamine (Dabeva et al., Am. J. Pathol. 143:1606, 1993). Efficacy of treatment can be determined by immunohistochemical staining for liver cell markers, microscopic determination of whether canalicular structures form in growing tissue, and the ability of the treatment to restore synthesis of liver-specific proteins. Liver cells can be used in therapy by direct administration, or as part of a bioassist device that provides temporary liver function while the subject's liver tissue regenerates itself following fulminant hepatic failure.

The efficacy of cardiomyocytes prepared according to this invention can be assessed in animal models for cardiac cryoinjury, which causes 55% of the left ventricular wall tissue to become scar tissue without treatment (Li et al., Ann. Thorac. Surg. 62:654, 1996; Sakai et al., Ann. Thorac. Surg. 8:2074, 1999, Sakai et al., J. Thorac. Cardiovasc. Surg. 118:715, 1999). Successful treatment will reduce the area of the scar, limit scar expansion, and improve heart function as determined by systolic, diastolic, and developed pressure. Cardiac injury can also be modeled using an embolization coil in the distal portion of the left anterior descending artery (Watanabe et al., Cell Transplant. 7:239, 1998), and efficacy of treatment can be evaluated by histology and cardiac function. Cardiomyocyte preparations embodied in this invention can be used in therapy to regenerate cardiac muscle and treat insufficient cardiac function (U.S. Patent 5,919,449 and WO 99/03973).

The examples that follow are provided by way of further illustration, and are not meant to imply any limitation in practicing the claimed invention.

EXAMPLES

Example 1: Feeder-free passage of hES cells

In this experiment, undifferentiated hES cells that had been maintained on primary mouse embryonic feeder cells were harvested, and then maintained in the absence of feeders. The culture wells were coated with Matrigel®, and the cells were cultured in the presence of conditioned nutrient medium obtained from a culture of irradiated primary fibroblasts.

Preparation of conditioned media (CM) from primary mouse embryonic fibroblasts (mEF):

Fibroblasts were harvested from T150 flasks by washing one time with $\text{Ca}^{++}/\text{Mg}^{++}$ free PBS and incubating in 1.5-2 mL trypsin/EDTA (Gibco) for about 5 min. After the fibroblasts detached from the flask, they were collected in mEF media (DMEM + 10% FBS). The cells were irradiated at 4000 rad (508 sec at 140kV: shelf setting 6 in a Torrex generator), counted and seeded at about 55,000 cells cm^{-2} in mEF media (525,000 cells/well of a 6 well plate). After at least 4 hours the media were exchanged with SR containing ES media (containing bFGF), using 3-4 mL per 9.6 cm^2 well of a 6 well plate. Conditioned media was collected daily for feeding of hES cultures. Alternatively, medium was prepared using mEF plated in culture flasks, exchanging medium daily at 0.3-0.4 mL cm^{-2} . Before addition to the hES cultures, the conditioned medium was supplemented with 4 ng/mL of human bFGF (Gibco). Fibroblast cultures were used in this system for about 1 week, before replacing with newly prepared cells.

Matrigel® coating:

Growth Factor Reduced Matrigel® or regular Matrigel® (Becton-Dickinson, Bedford MA) was thawed at 4°C. The Matrigel® was diluted 1:10 to 1:500 (typically 1:30) in cold KO DMEM. 0.75-1.0 mL of solution was added to each 9.6 cm^2 well, and incubated at room temp for 1 h. The coated wells were washed once with cold KO DMEM before adding cells. Plates were used within 2 h after coating, or stored in DMEM at 4°C and used within ~1 week.

Human ES culture:

Undifferentiated hES colonies were harvested from hES cultures on feeders as follows. Cultures were incubated in ~200 U/mL collagenase IV for about 5 minutes at 37 °C. Colonies were harvested by picking individual colonies up with a 20 μL pipet tip under a microscope or by scraping and dissociating into small clusters in conditioned medium (CM). These cells were then seeded onto Matrigel® in conditioned media at 15 colonies to each 9.6 cm^2 well (if 1 colony is ~10,000 cells, then the plating density is ~15,000 cells cm^{-2}).

The day after seeding on Matrigel®, hES cells were visible as small colonies (~100-2,000 cells) and there were single cells in-between the colonies that appeared to be differentiating or dying. As the hES cells proliferated, the colonies became quite large and very compact, representing the majority of surface area of the culture dish. The hES cells in the colonies had a high nucleus to cytoplasm ratio and had prominent nucleoli, similar to hES cells maintained on feeder cells. At confluence, the differentiated cells in between the colonies represented less than 10% of the cells in the culture.

Six days after seeding, the cultures had become almost confluent. The cultures were split by incubating with 1 mL ~200 U/mL Collagenase IV solution in KO DMEM for ~5 minutes at 37 °C. The collagenase solution was aspirated, 2 mL hES medium was added per well, and the hES cells were scraped from the dish with a pipette.

The cell suspension was transferred to a 15 mL conical tube, brought up to a volume of 6 mL, and gently triturated to dissociate the cells into small clusters of 10-2000 cells. The cells were then re-seeded on Matrigel® coated plates in CM, as above. Cells were seeded at a 1:3 or 1:6 ratio, approximately 90,000 to 170,000 cells cm⁻², making up the volume in each well to 3 mL. Medium was changed daily, and the cells were split and passaged again at 13 d and again at 19 d after initial seeding.

On day 19 after initial seeding, cells were harvested and evaluated for surface marker expression by immunofluorescence cell cytometry, using labeled antibodies specific for cell surface markers. For the hES cells maintained in the absence of feeders, a high percentage express SSEA-4, Tra-1-60 or Tra-1-81. These 3 markers are expressed on undifferentiated human ES cells that are maintained on feeders (Thomson et al., 1998). In addition, there is very little expression of SSEA-1, a glycolipid that is not expressed (or expressed at low levels) on undifferentiated ES cells. Immunohistochemical evaluation of SSEA-4, Tra-1-60 and Tra-1-81 indicates that the expression of these markers is localized to the ES colonies, not the differentiated cells in between the colonies.

Cultures of hES cells have been grown in the absence of feeder cells for over 180 days after initial seeding, with no apparent change in the proliferative capacity or phenotype. Human ES cells maintained on Matrigel® in mEF conditioned medium have a doubling time of about 31-33 hours, similar to the proliferation rate for hES cells grown on mEF feeder cells. H1 cells after 64 days of feeder-free culture showed a normal karyotype.

Example 2: Phenotypic markers of hES cells in feeder-free culture

Undifferentiated hES cells express SSEA-4, Tra-1-60, Tra-1-81, OCT-4, and hTERT. The expression of these markers decreases upon differentiation. In order to assess whether the cells maintained in feeder-free conditions retained these markers, cells were evaluated by immunostaining, reverse transcriptase PCR amplification, and assay for telomerase activity.

For analysis by fluorescence-activated cell sorting (FACS), the hES cells were dissociated in 0.5 mM EDTA in PBS and resuspended to about 5×10^5 cells in 50 µL diluent containing 0.1% BSA in PBS. For analyzing surface marker expression, cells were incubated in the primary antibodies, including IgG isotype control (0.5 µg/test), IgM isotype control (1:10), SSEA-1 (1:10), SSEA-4 (1:20), Tra-1-60 (1:40) and Tra-1-81 (1:80), diluted in the diluent at 4°C for 30 min. After washing with the diluent, cells were incubated with rat anti-mouse kappa chain antibodies conjugated with PE (Becton Dickinson, San Jose, CA) at 4°C for 30 min. Cells were washed and analyzed on FACScalibur™ Flow Cytometer (Becton Dickinson, San Jose, CA) using CellQuest™ software.

Similar to the hES cells on feeders, cells on Matrigel®, laminin, fibronectin or collagen IV expressed SSEA-4, Tra-1-60 and Tra-1-81. There was very little expression of SSEA-1, a glycolipid that is not expressed by undifferentiated hES cells.

For analysis by immunocytochemistry, cells were incubated with primary antibodies, including SSEA-1 (1:10), SSEA-4 (1:20), Tra-1-60 (1:40) and Tra-1-81 (1:80), diluted in knockout DMEM at 37°C for 30 min. Cells were then washed with warm knockout DMEM and fixed in 2% paraformaldehyde for 15 min. After washing with PBS, cells were incubated with 5% goat serum in PBS at room temp for 30 min, followed by incubation with the FITC-conjugated goat anti-mouse antibodies (1: 125) (Sigma) at room temp for 30 min. Cells were washed, stained with DAPI and mounted. The staining was typically performed ~2 days after passaging. Cells were also examined for expression of alkaline phosphatase, a marker for undifferentiated ES cells. This was performed by culturing the cells on chamber slides, fixing with 4 % paraformaldehyde for 15 min, and then washing with PBS. Cells were then incubated with alkaline phosphatase substrate (Vector Laboratories, Inc., Burlingame, CA) at room temperature in the dark for 1 h. Slides were rinsed for 2-5 min in 100% ethanol before mounting.

The results showed that SSEA-4, Tra-1-60, Tra-1-81, and alkaline phosphatase were expressed by the hES colonies on Matrigel® or laminin, as seen for the cells on feeders — but not by the differentiated cells in between the colonies.

Figure 1 shows OCT-4 and hTERT expression of H1 cells on feeders and off feeders, as detected by reverse-transcriptase PCR amplification. For radioactive relative quantification of individual gene products, QuantumRNA™ Alternate18S Internal Standard primers (Ambion, Austin TX, USA) were employed according to the manufacturer's instructions. Briefly, the linear range of amplification of a particular primer pair was determined, then coamplified with the appropriate mixture of alternate18S primers:competimers to yield PCR products with coinciding linear ranges. Before addition of AmpliTaq™ (Roche) to PCR reactions, the enzyme was pre-incubated with the TaqStart™ antibody (ProMega) according to manufacturer's instructions. Radioactive PCR reactions were analyzed on 5% non-denaturing polyacrylamide gels, dried, and exposed to phosphoimage screens (Molecular Dynamics) for 1 hour. Screens were scanned with a Molecular Dynamics Storm 860 and band intensities were quantified using ImageQuant™ software. Results are expressed as the ratio of radioactivity incorporated into the hTERT or OCT-4 band, standardized to the radioactivity incorporated into the 18s band.

Primers and amplification conditions for particular markers are as follows. OCT-4: Sense (SEQ. ID NO:2) 5'-CTTGCTGCAG AAGTGGGTGG AGGAA-3' AntiSense (SEQ. ID NO:3) 5'-CTGCAGTGTG GGTTCGGGC A-3'; alternate18:competimers 1:4; 19 cycles (94° 30 sec; 60° 30 sec; 72° 30 sec). hTERT: Sense (SEQ. ID NO:4) 5'-CGGAAGAGTG TCTGGAGCAA-3' AntiSense (SEQ. ID NO:5) 5'-GGATGAAGCG GAGTCTGGA-3'; alternate18:competimers 1:12; 34 cycles (94° 30 sec; 60° 30 sec; 72° 30 sec).

hTERT and OCT-4 expression was seen in all the culture conditions except Matrigel® and regular medium. Furthermore, after exposure of cells to retinoic acid (RA) or dimethyl sulfoxide (DMSO), factors that promote cell differentiation, the expression of hTERT was markedly decreased.

Figure 2 shows telomerase activity measured by TRAP activity assay (Kim et al., Science 266:2011, 1997; Weinrich et al., Nature Genetics 17:498, 1997). All the cultures conditions showed positive telomerase activity after 40 days on Matrigel®, laminin, fibronectin or collagen IV in mEF conditioned medium.

Example 3: Differentiation of hES cells

In this experiment, differentiation using standard methods of aggregate formation was compared with a direct differentiation technique.

For the aggregate differentiation technique, monolayer cultures of rhesus and human ES lines were harvested by incubating in Collagenase IV for 5-20 min, and the cells were scraped from the plate. The cells were then dissociated and plated in non-adherent cell culture plates in FBS-containing medium. The plates were placed into a 37°C incubator, and in some instances, a rocker was used to facilitate maintaining aggregates in suspension. After 4-8 days in suspension, aggregate bodies formed and were plated onto a substrate to allow for further differentiation.

For the direct differentiation technique, suspensions of rhesus and human ES cells were prepared in a similar fashion. The cells were then dissociated by trituration to clusters of ~50-100 cells, and plated onto glass coverslips treated with poly-ornithine. The cells were maintained in serum containing medium, or defined medium for 7-10 days before analysis.

Cells from both preparations were fixed and tested by immunoreactivity for β -tubulin III and MAP-2, which is characteristic of neurons, and glial fibrillary acidic protein (GFAP), which is characteristic of astrocytes. Results are shown in Table 1.

TABLE 1: Comparison of hPS Differentiation Methods

ES Cell Line used for differentiation		Differentiation via Aggregate Bodies		Direct Differentiation	
		Neurons	Astrocytes	Neurons	Astrocytes
R366.4	(Rhesus line)	+	+	+	+
R278.5	(Rhesus line)	+	+	+	+
R456	(Rhesus line)	+	+	+	+
H9	(Human line)	+	+	+	+
H9.1	(Clone of H9)	(Not Done)	(Not Done)	+	+
H9.2	(Clone of H9)	+	+	+	+

Rhesus and human ES lines differentiated into cells bearing markers for neurons and astrocytes, using either the aggregate or direct differentiation technique. In the rhesus cultures, percentage of aggregates that contained neurons ranged from 49% to 93%. In the human lines examined, the percentage of aggregates containing neurons ranged from 60% to 80%. Double labeling for GABA and β -tubulin indicated that a sub-population of the neurons express the inhibitory neurotransmitter GABA. In addition, astrocytes and oligodendrocytes were identified with GFAP immune reactivity and GalC immune reactivity, respectively. Therefore, the human and rhesus ES cells have the capacity to form all three major cell phenotypes in the central nervous system.

The effect of several members of the neurotrophin growth factor family was examined. hES cells were differentiated by harvesting with collagenase, dissociating, and reseeding onto poly-ornithine coated cover slips. The cells were plated into DMEM/F12 + N2 + 10% FBS overnight. The following day, the serum was removed from the medium and replaced with 10 ng/mL human bFGF and the growth factor being tested. After 24 hours, bFGF was removed from the medium. These cultures were fed every other day. They were fixed after 7 days of differentiation and immunostained for analysis. The number of neurons was evaluated by counting cells positive for β -tubulin. Cultures maintained in the presence of 10 ng/mL brain derived neurotrophic factor (BDNF) formed approximately 3-fold more neurons than the control cultures. Cultures maintained in neurotrophin-3 (1 ng/mL) formed approximately 2-fold more neurons than control cultures.

In a subsequent experiment, suspensions of human ES cells were prepared from parental line H9 and two subcloned lines. The cells were harvested using collagenase IV, and then replated onto poly-ornithine coated glass slides in medium containing 20% FBS. The cultures were then fed every other day for 7-10 days, then fixed for immunostaining. From each of these lines, a number of differentiated cells stained positively for muscle-specific actin (antibody from Dako), but were negative for cardiac troponin I. Several patches of cells stained positively for α -fetoprotein, indicating the presence of endoderm cells.

Example 4: Comparison of direct differentiation with differentiation through embryoid bodies

To induce direct differentiation, undifferentiated hES cells were harvested and re-plated directly into differentiating conditions. Considerable cell death was apparent upon plating, but many cells adhered and began to proliferate and/or differentiate. In cultures differentiated using serum containing conditions, the cultures continued to proliferate and reached confluence within 5-10 days. At this time, the cultures contained a heterogeneous population that displayed many different morphologies. Immunocytochemistry revealed ectoderm, mesoderm and endoderm lineages using antibodies against β -tubulin III, muscle specific actin and α -fetoprotein, respectively. The positive staining for all of these cell types appeared in patches that were sometimes quite dense, therefore it was difficult to accurately quantify the percentages of each cell type.

In order to increase the percentage of neurons, the hES cells were plated onto poly-ornithine coated glass coverslips and cultures in defined media. Although these data indicate that cells from all three germ layers can be derived without the production of EBs, cardiomyocytes were not identified.

By way of comparison, hES cells were induced to differentiate by generating embryoid bodies (EBs). In these experiments, ES cells were harvested and replated in suspension cultures. Although initially a marked amount of cell death was observed, after 2-3 days the remaining cells formed aggregates. EBs were maintained for as many as 16 days in culture and were still viable and formed many structures after subsequent plating. Later stage human EBs often showed a cystic morphology and sometimes gave rise to beating EBs.

To assess cardiomyocyte formation, EBs were transferred to gelatin-coated plates or chamber slides after 4 days in the suspension cultures. The EBs attached to the surface after seeding, proliferated and differentiated into different types of cells. Spontaneously contracting cells were observed in various regions of the culture at differentiation day 8 and the number of beating regions increased until about day 10. In some cases, more than 75% of the EBs had contracting regions. Beating cells were morphologically similar to mouse ES cell-derived beating cardiomyocytes. In addition, the expression of the cardiac specific marker cardiac troponin I was examined at differentiation day 15 using immunocytochemistry. Individual contracting foci in the differentiated cultures were photographed to record the contracting area before the culture was fixed. The culture was then evaluated for cardiac cTnI expression and matched to the original photographs to determine the percentage of contracting areas that were positive for cTnI staining. As a control, cells adjacent to the contracting foci were also examined for cTnI staining. In these cultures 100% of the contracting areas showed positive immunoreactivity, while minimal immunoreactivity was observed in the non-beating cells.

Cultures of differentiated EBs were subjected to Western blot analysis using monoclonal antibody against cTnI. This assay gave a strong 31 kDa protein signal, corresponding to the size of the purified native human cardiac Tn I. cTnI was detected in differentiated human ES cells containing contracting cells but not in undifferentiated ES cells or differentiated cultures with no evidence of contracting cells, suggesting the specific detection of cardiomyocytes. As a control, the blot was reprobed with β -actin specific antibody, confirming the presence of similar amounts of proteins in all samples.

In other experiments, EBs were cultured for 8 or 16 days and maintained as adherent cultures for an additional 10 days. RNA was prepared from the differentiated human ES cells and semiquantitative RT-PCR was performed to detect the relative expression of the endoderm-specific products α_1 -anti-trypsin, AFP, and albumin. Low levels of α_1 -anti-trypsin and AFP were detected in the undifferentiated cultures; little or no albumin was detected in the same cultures. All 3 markers were detected at significantly higher levels after differentiation. Expression of all 3 endoderm markers was higher in cultures derived from 8 day embryoid bodies than 16 day embryoid bodies.

Example 5: Transfection and transduction of hES cells maintained on primary mEF feeder layers

hES cultures were maintained in a growth medium composed of 80% KO DMEM (Gibco) and 20% Serum Replacement (Gibco) supplemented with 1% non-essential amino acids, 1 mM glutamine, 0.1 mM β -mercaptoethanol and 4 ng/mL hbFGF (Gibco).

Plates were coated with a solution of 0.5% gelatin (Sigma) at 37° overnight prior to the addition of cells. Primary mEFs were cultured in standard mEF medium, and split 1:2 every 2 days for up to 5 splits. Subconfluent cultures of mEFs were detached with trypsin, resuspended in 10 mL medium, and irradiated with a cumulative dose of 3500-4000 rads with a Torrex 150D X-ray generator. Irradiated cells were pelleted at $400 \times g$ for 5 min and resuspended at 1.25×10^5 cells per mL in standard mEF medium. Individual wells of a 6-well plate were seeded with 3.75×10^5 irradiated mEFs per well; individual wells of a 24-well plate were seeded with 75,000 irradiated mEFs per well.

Transfection was performed as follows. hES cells plated in 6 well plates were removed from the feeder layer with collagenase (~200 units/mL) at 37° for 7-10 min. When colonies began to detach, the collagenase from each well was aspirated and replaced with 2 mL of standard hES growth medium/well. The hES cells were removed by scraping the surface of a single well with a 5 mL pipet and transferred to a 50 mL conical tube. Additional hES growth medium was added to a final volume of 10 mL. The cell suspension was triturated 10-12 times with a 10 mL pipet, and an additional 8 mL of standard hES growth medium added. Three mL of the cell suspension were added to each well of 6 well plates that were pre-coated with gelatin and mEF feeder layers as described above (i.e., 1 well of a 6 well plate was sufficient to seed 6 wells of a new plate).

Replated hES cells were tested with a number of different transfection systems to determine whether genetic alteration of hES cells could be achieved without causing differentiation. Systems tested included the following: Mammalian Transfection Kit (CaPO4 and DEAE reagents), Stratagene cat # 200285; TransIT-LT1 Mirus (Panvera), cat # MIR 2310; Polybrene (Sigma); Poly-L-Lysine (Sigma); Superfect™ (Qiagen); Effectene™ (Qiagen); Lipofectin™ (Life Technologies); Lipofectamine (differs from Lipofectamine 2000™) (Life Technologies); Cellfectin™ (Life Technologies); DMRIE-C (Life Technologies); Lipofectamine 2000 (Life Technologies); and electroporation using BioRad™ Gene pulser.

Under the conditions used, Lipofectamine 2000™ (Gibco Life Technologies cat # 11668019, patent pending) and FuGENE™ (trademark of Fugent L.L.C.; a proprietary blend of lipids and other components, purchased from Roche Diagnostic Corporation cat # 1 814 443) both resulted in good transfection efficiency. The efficiency was generally best if these reagents were contacted with replated hES cells ~48 h after the replating.

Transfection using Lipofectamine 2000™ was conducted as follows: The plasmid DNA (3-5 μ g of pEGFP-C1, ClonTech cat. # 6084-1) was diluted in water to a final volume of 100 μ L. In pilot experiments, 5 to 30 μ L of Lipofectamine 2000™ (Gibco, cat # 11668-019) were diluted in OptiMEM™ (Gibco, cat # 11-58-021) to a final volume of 100 μ L. The DNA solution was then added slowly to the Lipofectamine2000™ solution and mixed gently. The mixture was incubated at room temperature for 20-30 min before being supplemented with 800 μ L of OptiMEM™. Cells were washed with 3 mL of pre-warmed OptiMEM™ and incubated in 0.5-1 mL of the DNA/lipid mixture solution at 37°C for 4 h, per well (9.6 cm²). In some experiments, after 4 h, the complex was removed before the addition of 4 mL of mEF-conditioned medium; in others, sufficient mEF-conditioned medium was added to the wells to reach a final volume of 3.5 mL and the mixture was left on the cells overnight. In other experiments the DNA/lipid mixture was added to wells containing sufficient mEF-conditioned medium such that the final volume was 3.5 mL, and the cells were incubated in this mixture overnight.

Transfection using FuGENE™ was conducted as follows. Each well was transfected with 10 µg DNA using FuGENE™ 6 (Roche Diagnostics Corp.), at a ratio of 3:2 FuGENE™ reagent to DNA as described by the manufacturer's directions. OptiMEM™ serum-free medium was used in the transfections. In the "old protocol", 4 h after the addition of the FuGENE™-DNA complex, 2.5 mL of standard hES growth medium was added to each transfected well. In the revised protocol ("3:2 L"), transfected wells were not re-fed with standard hES growth medium. Twenty-four hours after transfection, GFP-expression was assessed by flow cytometry.

Forty-eight hours before transfection, hES cells were seeded onto 6 well plates that had been coated with gelatin and mEF feeder layers as described above. hES cells were transfected using FuGENE™ 6 (Roche) or Lipofectamine 2000™ (Gibco) according to the manufacturers' instructions. Twenty-four hours after transfection, cells were assessed for GFP expression by inspection under a fluorescent microscope or flow cytometry. In the experiment shown in Fig. 1, three methods were compared: the standard Lipofectamine 2000™ protocol, the standard FuGENE™ protocol, and a variant FuGENE™ protocol in which the DNA/lipid mix was left on the cells overnight. The results demonstrated that while Lipofectamine 2000™ consistently yielded a higher percentage of GFP-expressing cells, the variant FuGENE™ protocol resulted in GFP-expressing cells with a higher mean fluorescence intensity.

Transient transductions using adenoviral vectors were conducted as follows. The vector Ad5CMV5-GFP (referred to here as Ad5GFP) contains the green fluorescent protein encoding region under control of the CMV promoter, and was purchased from Quantum Biotechnologies, cat # ADV0030. Seventy-two hours before transduction, hES cells were seeded onto 24 well plates that had been coated with gelatin and mEF feeder layers as described above. Before transduction, 3 wells of hES cells were detached with a solution of 0.05% trypsin/5mM EDTA (Sigma) at 37°, resuspended in 500 µL of standard mEF growth medium, and counted with a hemocytometer (the 75,000 mEF feeder cells were subtracted from each well) to establish the cell number before transfection. The adenovirus stock was thawed on ice immediately prior to use.

For infection with Ad5GFP, growth media was aspirated from the wells containing hES cells and replaced with 1 mL of hES growth medium plus 9 µL of Ad5 GFP stock (MOI of 40). Two hours later, the virus-containing medium was replaced with 1 mL of hES growth medium per well. Each transduced well was re-fed with 1 mL of fresh hES growth medium every 24 hours. GFP expression was assessed by flow cytometry. The results from a typical experiment indicated that expression was highest at 24 hr after transduction but persisted for at least 8 days at low levels (by the later time points, extensive differentiation had occurred due to overgrowth of the hES cells).

Example 6: Preparation of the immortalized feeder cell line NH190

In this example, a permanent mouse cell line was established that is suitable for conditioning medium for the culture of primate pluripotent stem (pPS) cells. The NHG190 line is a mouse embryonic fibroblast cell line immortalized with telomerase that is triple drug resistant, and expresses green fluorescent protein (GFP).

Two mouse strains were obtained from Jackson Laboratory (Bar Harbor, Maine) that have a transgene for resistance to the antibiotics neomycin or hygromycin. The C57BL/6J TgN(pPGKneobpA)3Ems mice and C57BL/6J-TgN(pPWL512hyg)1Ems mice from Jackson Labs were cross-bred. Embryos that were both neomycin- and hygromycin-resistant were dissected at day 13.5 post conception according to standard protocols for preparing mouse embryonic fibroblasts (mEF) for feeder layers (E.J. Robertson, pp. 71-112 in *Teratocarcinoma and Embryonic Stem Cell Lines*, ed. E. J. Robertson, Oxford: IRL Press, 1987). The derived mEF cells were stored frozen.

The mEFs were thawed in growth medium containing 20% fetal calf serum (HyClone), 2 mM L-glutamine (Gibco/BRL), 80% DMEM (Gibco/BRL). The cells were expanded using 1:2 split ratios for 4 passages. Two flasks that had reached ~75% confluence were fed with fresh medium 4 h before electroporation. Cells were removed from the flasks with 0.5% trypsin/500 mM EDTA (Gibco/BRL), pelleted at $400 \times g$ for 5 min at room temperature, and resuspended in the growth medium at a concentration of 4×10^6 cells/mL.

The cell suspension was divided into two 500 μ L aliquots and transferred to two 0.4 cm gap electroporation cuvettes (BioRad). One cuvette received 5 μ g of the control plasmid (pBS212; puromycin-resistance gene driven by the SV40 early enhancer/promoter); the other received 5 μ g of pGRN190, comprising the murine telomerase reverse transcriptase (mTERT) coding region driven by MPSV promoter plus puromycin resistance gene driven by the SV40 early enhancer/promoter. The cells and DNA were mixed by hand, and electroporated using a BioRad gene Pulser with a BioRad capacitance extender at a setting of 300V, 960 μ F.

Each aliquot of cells was transferred to an individual 150 cm plate containing 25 mL of growth medium. The medium on the plates was exchanged on the following day, and on the next day, growth medium was replaced by growth medium plus 0.5 μ g/mL puromycin. The medium on the plates was exchanged for fresh puromycin-containing medium every 48 hrs until 29 days after electroporation. At this time, large individual colonies of puromycin-resistant cells were evident in both the pBS212- and pGRN190- electroporated plates. Ten colonies from the control plate and 12 from the pGRN190-electroporated plate were isolated with cloning cylinders and each colony was transferred to 1 well of a 48-well plate (1 well per colony).

One week later, all surviving colonies that had expanded to reach confluence in the 48 well plate (three control colonies, 1 pGRN190-electroporated colony) were transferred individually to wells of a 24 well plate. Six days later, the only colony that had continued to expand was derived from the pGRN190-electroporated plate, and was subsequently designated **NH190**. The cells were maintained in growth medium plus 0.5 μ g/mL puromycin. Analysis for telomerase activity by TRAP assay (Kim et al., Nucleic Acids Res. 25:2595, 1997) demonstrated that NH190 cells express functional telomerase activity.

To facilitate monitoring of the cells in mixed culture populations and in vivo, NH190 cells were further infected with a retroviral construct conferring expression of green fluorescent protein (GFP). The enhanced GFP sequence from plasmid pEGFP-1 is one of the Living Colors™ fluorescent protein vectors, available from ClonTech. It contains an enhanced GFP encoding region, with changes that alter restriction nuclease cleavage sites, and shift the excitation and emission wavelengths of the encoded protein. The EGFP-1 sequence was cloned into the vector pMSCV.neo, ClonTech cat # K1062-1. NH190 cells were transduced with the engineered vector, and GFP positive cells were separated by FACS sorting. The GFP expressing cell line was designated **NHG190**. These cells have been carried in culture for over 3 months.

Example 7: Genetic modification of hES cells maintained on NHG190 feeder cells

NHG190 cells were cultured in DMEM (Gibco) plus 20% fetal bovine serum (HyClone) and 5 mM glutamine. Cells were split 1:10 every 3 d. Subconfluent cultures were detached with trypsin, suspended in 10 mL medium, and irradiated with a cumulative dose of 3500 rads with a Torrex 150D X-ray generator. Irradiated cells were pelleted at $400 \times g$ for 5 min and resuspended at 1.25×10^5 cells per mL in either NHG190 medium or standard hES medium.

Conditioned medium was prepared by plating NHG190 cells at 4.08×10^4 cm⁻¹ on gelatin-coated plates. At 18-24 h after plating, medium was exchanged for standard hES medium with 4 ng/mL added bFGF. The medium was conditioned by the cells for 18-24 h, harvested, and an additional 4 ng/mL bFGF was added. The

medium was used to support hES cell cultures the same day as it was collected. Irradiated NHG190 cells could be used for preparing conditioned medium for 7-10 days.

hES cells were transfected as follows. The cells were removed from the feeder layer using collagenase (~200 U/mL) at 37°C for 7-10 min, and transferred to a 50 mL conical tube. hES growth medium was added to a final volume of 10 mL; the suspension was triturated 10-12 times with a 10 mL pipet, and another 8 mL hES medium was added. Three mL of cell suspension was added to each well in a 6-well plate precoated with Matrigel® and NHG190 feeder cells.

Forty-eight hours after seeding, the hES were transfected with 10 µg DNA per well using FuGENE™ 6 (Roche) according to manufacturer's protocol in OptiMEM™ serum-free medium. The DNA was a plasmid containing the PGK promoter driving neo^r. Four h later, 3 mL of NHG190-conditioned medium was added to each transfected well. Cells were re-fed daily with 3 mL conditioned medium. Forty-eight h after transfection, the cells were layered with NHG190 conditioned medium containing 200 µg/mL added geneticin (Sigma), which was replaced daily thereafter. After 3 days of selection, additional irradiated NHG190 feeder cells were added (1.25 × 10⁵ cells/well in hES medium). Twenty-four h later, the medium was again replaced with NHG190-conditioned medium containing 200 µg/mL geneticin, replaced daily.

Individual colonies were isolated and expanded through another round of selection. After a further 5 days, individual colonies were identified by microscope and marked on the outside of the dish. Medium was removed, and replaced with collagenase (~200 U/mL). Individual colonies were picked using a p20 pipet tip, and transferred to individual tubes containing 2 mL NHG190 conditioned medium (without geneticin). The suspension was triturated 5 times to disaggregate colonies, and the contents of each tube were transferred to a well of a 12-well plate coated with gelatin and irradiated NHG190 cells (1.875 × 10⁵ cells/well). Cells were fed 24 h later with 2 mL fresh conditioned medium. Two days after seeding, cells were layered with 2 mL conditioned medium containing 200 µg/mL geneticin, replaced daily for 5 days. As each well became 50-75% confluent, the cells were detached with collagenase, transferred to 6 mL conditioned medium, and triturated 10-12 times. 3 mL cell suspension was added to each of 2 wells of a 6-well plated coated with gelatin and irradiated NHG190 cells (3.75 × 10⁵ cells/well); the cells were re-fed with 3 mL conditioned medium at 24 h. The cells were then selected for 5 days using 3 mL conditioned medium containing geneticin, and split 1:6 as before.

Stable transduction using retrovirus was conducted as follows. Retroviral vector designated GRN354 was constructed at Geron Corp. using PMSCVneo vector purchased from ClonTech (cat # K1062-1). The eGFP encoding region was inserted downstream from the MSCV LTR. The LTR drives expression of GFP and the vector also contains the neo^r gene driven by the murine PGK promoter. Plates were coated with 0.5% gelatin and NHG190 feeder cells (7.5 × 10⁴ in 1 mL NHG190 medium for 24 well plates; 3.75 × 10⁵ in 3 mL medium for 6 well plates). The hES line H7 was seeded onto a 24 well prepared plate in hES medium (1 mL/well). Forty-eight h later, 3 wells of hES cells were detached using 0.05% trypsin/5 mM EDTA (Sigma) at 37°C, resuspended in 500 µL NHG190 medium, and counted. Stock of retrovirus construct pGRN354 was thawed on ice immediately prior to use. Growth medium was aspirated from the wells and replaced with 400 µL hES medium plus 8 µL retrovirus (MOI of 10) and 4 µL of 8 mg/mL polybrene solution (Sigma). Two h later, 800 µL hES growth medium were added per well. Each transduced well was re-fed with 1 mL fresh hES medium every 24 h.

Four days after transduction, medium was replaced with 1 mL hES growth medium containing 200 µg/mL geneticin. After 3 days of geneticin selection, the cells were detached with collagenase, triturated, resuspended in 3 mL hES medium, reseeded into one well of a 6-well plate coated with gelatin and NHG190 feeders, and re-fed with hES medium after 24 h. The medium was then again replaced with hES medium containing geneticin and

refed every 24 h. Undifferentiated colonies survived the selection, and have been maintained for over 3 months. FACS analysis showed that 50-65% of the selected cells express GFP, albeit at low levels. The karyotype of the cells was normal.

Figur 3 shows GFP expression of hES cells transduced with retrovirus and then differentiated. The hES cell line H7 was plated on drug-resistant (NHG190) feeder layers, infected with GRN354 and selected for resistance to the drug G418. Transduced cells were expanded and maintained under G418 selection for multiple passages. The cells were transferred to suspension culture to form embryoid bodies, allowed to differentiate for 4 days, and then plated in 20% FBS medium for 1 week. After extensive differentiation occurred, cultures were fixed in 4% paraformaldehyde and photographed under fluorescence for GFP expression. Many of the differentiated cells express higher levels of GFP than the undifferentiated transfected hES line, consistent with differential activation of the MESV-LTR in different cell types.

Example 8: Transfection of feeder-free hES cells

In this example, hES cells maintained in feeder-free culture on laminin in conditioned medium were genetically modified by transfecting with a plasmid carrying green fluorescent protein (GFP) driven by the CMV promoter.

mEF conditioned medium was prepared as described earlier. mEFs were irradiated and seeded at about 5.7×10^4 cells/cm². After at least 16 hours the medium was exchanged with hES medium including 4 ng/mL added hbFGF. Conditioned medium was collected daily for feeding of hES cultures. Before addition to the hES cultures this medium was supplemented with an additional 4 ng/mL of hbFGF. Where needed for selection of stable transfectants, the mEF-conditioned medium was supplemented with 200 µg/mL geneticin (Sigma cat. # G5013).

H9 hES cells maintained on mEF feeder layers were harvested from cultures by incubation with ~200 units/mL collagenase IV at 37°C for 10 min. Cells were dissociated and resuspended in regular hES culture medium or mEF-conditioned medium. Cells in the regular medium were then re-seeded onto mEF feeder layers and cells in the mEF-conditioned medium were plated onto Matrigel® or laminin. Seeding density for all cultures was approximately 4×10^4 cells/cm². Cells on feeder layers were maintained in regular medium while cells on matrices were maintained in mEF- conditioned medium for 1 or 2 days before the transfection. Conditioned medium was replaced every 24 h.

hES cell cultures were transfected with Lipofectamine 2000™ as described above. FACS analysis of GFP expression was conducted as follows. hES cells were harvested using 0.5mM EDTA in PBS and resuspended at approximately 1×10^6 cells/test. Cells were washed in a solution containing PBS plus 2% FBS, 0.1% sodium azide, and 2 mM EDTA. SSEA-4 staining was performed in the same buffer using antibody obtained from the Developmental Studies Hybridoma Bank (University of Iowa, Iowa City) at 1:15 dilution. Isotype matched controls were obtained from Sigma, (St. Louis MO, USA). Cells were incubated with antibodies in a final volume of 100 µl for 30 min at 4°C, washed and incubated with rat anti-mouse κ chain antibodies conjugated with PE (Becton Dickinson, San Jose, CA) at 4°C for 30 min. Samples were washed as before and analyzed for GFP and SSEA-4 expression on FACScalibur™ flow cytometer (Becton Dickinson, San Jose, CA) using CellQuest™ software.

hES cells of the H9 line maintained on laminin in mEF-conditioned medium were transfected with a plasmid carrying GFP driven by the CMV promoter at 24 or 48 h after plating. Initial experiments used a mixture of 5 µg of plasmid and 12 µL of Lipofectamine 2000™. Cells received 1 mL of DNA/lipid complex and were incubated for 4 h at 37° before the addition of 3 mL of mEF-conditioned medium, and then monitored for GFP expression 24 h after transfection.

Figure 4 shows the results of this experiment. Panel A: morphology of H9 cells maintained on laminin. Panel B: GFP-positive cells observed in the same colony shown in A. Panel C: FACS analysis of % GFP-positive cells in SSEA-4 high population(undifferentiated cells). Cells were transfected 24 (bar 1 and 2) or 48 h (bar 3 and 4) after the seeding and analyzed 24 (bar 1 and 3) or 48 h (bar 2 and 4) after the transfection. Bright green cells were observed in compact areas of undifferentiated ES colonies on laminin 24 h after transfection (Panels A & B). Transfection at 48 h after initial seeding gave the highest efficiency: 38% of the cells were GFP-positive as determined by FACS analysis 24 h after the transfection (Panel C).

The next experiment compared the transfection efficiency of H9 cells maintained on Matrigel® or laminin-coated plates in mEF-conditioned medium with cells maintained on mEF feeders. Cells on feeder layers maintained in regular medium were used as a control. Morphological differences between cells on feeders and cells off feeders were observed 1 or 2 days after seeding. Colonies on feeders were more compact than cells maintained off feeder layers; individual hES cells in feeder-free cultures were less compact and flatter. There was no significant difference in cell or colony morphology between cells on laminin and cells on Matrigel. These cells were transfected with a plasmid expressing GFP driven by the CMV promoter 2 days after seeding. Twenty-four hours after the transfection, cells were examined for GFP expression under a fluorescence microscope.

The cells were maintained on mEF feeders in regular medium (mEF/RM), on laminin in medium conditioned by mEF (Laminin/CM) or on Matrigel® in the conditioned medium (Matrigel/CM). Bright green cells were observed in undifferentiated hES colonies of feeder-free cultures. In contrast, very few green cells were found in colonies on feeders. FACS analysis showed that 16% of cells on Matrigel® and 14% of cells on laminin were GFP positive in SSEA-4 high population while only 5% of cells on feeders were positive. These results indicate that transfection efficiency is significantly increased by using feeder-free conditions.

The next experiments evaluated the effects of 1) the ratio of DNA:lipid; 2) adding the DNA/lipid complex to cells 4 h prior to the addition of mEF-conditioned medium vs. addition of the complex to cells in the presence of mEF-conditioned medium; and 3) use of Lipofectamine 2000™ vs. FuGENE™.

Transfection using Lipofectamine2000™ is described above. Transfection with FuGENE™ was conducted as follows. The plasmid DNA (5-10 µg of pEGFP-C1, ClonTech cat. # 6084-1) was diluted in water to a final volume of 100 µl. In pilot experiments, 5-30 µL of FuGENE™ were added to sufficient OptiMEM™ to achieve a final volume of 100 µL. The DNA solution was then added slowly to the FuGENE™ solution and mixed gently. The mixture was incubated at room temperature for 30 min before being supplemented with 800 µl of OptiMEM™. Cells were washed with 3 mL of pre-warmed OptiMEM™ and incubated in 1 mL of the DNA/lipid mixture solution at 37°C for 4 h. In some experiments, after 4 h the wells received an additional 2 mL of mEF-conditioned medium; in others the DNA/lipid mixture was added to wells containing 2 mL of mEF-conditioned medium and the cells were incubated in this mixture overnight.

Highest efficiencies were obtained under the following conditions: Bar 1 = a mixture of 5 µg plasmid plus 12 µl of Lipofectamine 2000™, adding 1 mL of the DNA/lipid mixture to wells containing 2.5 mL of mEF-conditioned medium and incubating the cells in this mixture overnight. Bars 2 & 3 = a mixture of 10 µg plasmid plus 15 µl of FuGENE™ and incubating the cells in 1 mL of the DNA/lipid mixture for 4 h before adding 2.5 mL of mEF-conditioned medium. L=Lipofectamine2000™; F=FuGENE™.

To investigate whether the feeder-free hES cells undergo stable genetic modification, H1 hES cells maintained on Matrigel® were cotransfected with a mixture of 7.5 µg plasmid carrying β-galactosidase driven by the EF1a promoter, and 2.5 µg of plasmid carrying the PGK promoter driving the neophosphotransferase gene. The cells were transfected 48 h after plating them on Matrigel® in mEF-conditioned medium. 10 µg of plasmid

plus 15 μ l of FuGENE™ were incubated with the cells in 1 mL for 4 h before adding 2.5 mL of mEF-conditioned medium. After 48 h, medium was exchanged for mEF-conditioned medium supplemented with 200 μ g/mL geneticin. Cultures were maintained in this geneticin-containing medium with daily medium exchange for over 21 days. All mock-transfected cultures (i.e., those that received FuGENE™ mixed with water rather than plasmid) died within 48-72 h. Drug resistant colonies arose in the wells transfected with both FuGENE™ and plasmid at a frequency of about 1 in to 10⁵ originally transfected cells. The colonies were maintained in geneticin-containing mEF-conditioned medium and expanded.

Example 9: Preparation of vectors in which a thymidine kinase gene is under control of an hTERT promoter sequence

The lambda clone designated λ G Φ 5 containing the hTERT promoter is deposited with the American Type Culture Collection (ATCC), 10801 University Blvd., Manassas, VI 20110 U.S.A., under Accession No. 98505. λ G Φ 5 contains a 15.3 kbp insert including approximately 13,500 bases upstream from the hTERT coding sequence.

A Not1 fragment containing the hTERT promoter sequences was subcloned into the Not1 site of pUC derived plasmid, which was designated pGRN142. A subclone (plasmid "pGRN140") containing a 9 kb NcoI fragment (with hTERT gene sequence and about 4 to 5 kb of lambda vector sequence) was partially sequenced to determine the orientation of the insert. pGRN140 was digested using Sall to remove lambda vector sequences, the resulting plasmid (with removed lambda sequences) designated pGRN144. The pGRN144 insert was then sequenced.

SEQ. ID NO:1 is a listing of the sequence data obtained. Nucleotides 1-43 and 15376-15418 are plasmid sequence. Thus, the genomic insert begins at residue 44 and ends at residue 15375. The beginning of the cloned cDNA fragment corresponds to residue 13490. There are Alu sequence elements located ~1700 base pairs upstream. The sequence of the hTERT insert of pGRN142 can now be obtained from GenBank (<http://www.ncbi.nlm.nih.gov/>) under Accession PGRN142.INS AF121948. Numbering of hTERT residues for plasmids in the following description begins from the translation initiation codon, according to standard practice in the field. The hTERT ATG codon (the translation initiation site) begins at residue 13545 of SEQ. ID NO:1. Thus, position -1, the first upstream residue, corresponds to nucleotide 13544 in SEQ. ID NO:1.

Expression studies were conducted with reporter constructs comprising various hTERT upstream and intron sequences. A BglII-Eco47III fragment from pGRN144 (described above) was digested and cloned into the BglII-NruI site of pSEAP2Basic (ClonTech, San Diego, CA) to produce plasmid designated pGRN148. A second reporter-promoter, plasmid pGRN150 was made by inserting the BglII-FspI fragment from pGRN144 into the BglII-NruI sites of pSEAP2. Plasmid pGRN173 was constructed by using the EcoRV-StuI (from +445 to -2482) fragment from pGRN144. This makes a promoter reporter plasmid that contains the promoter region of hTERT from approximately 2.5 kb upstream from the start of the hTERT open reading frame to just after the first intron within the coding region, with the initiating Met codon of the hTERT open reading frame changed to Leu. Plasmid pGRN175 was made by APA1(Klenow blunt)-SRF1 digestion and religation of pGRN150 to delete most of the Genomic sequence upstream of hTERT. This makes a promoter/reporter plasmid that uses 204 nucleotides of hTERT upstream sequences (from position -36 to -117). Plasmid pGRN176 was made by PML1-SRF1 religation of pGRN150 to delete most of the hTERT upstream sequences. This makes a promoter/reporter plasmid that uses 204 nucleotides of hTERT upstream sequences (from position -36 to -239).

Levels of secreted placental alkaline phosphatase (SEAP) activity were detected using the chemiluminescent substrate CSPDTM (ClonTech). SEAP activity detected in the culture medium was found to be directly proportional to changes in intracellular concentrations of SEAP mRNA. The pGRN148 and pGRN150 plasmids (hTERT promoter-reporter) and the pSEAP2 plasmid (positive control, containing the SV40 early promoter and enhancer) were transfected into test cell lines. pGRN148 and pGRN150 constructs drove SEAP expression as efficiently as the pSEAP2 in immortal (tumor-derived) cell lines. Only the pSEAP2 control gave detectable activity in mortal cells.

The ability of the hTERT promoter to specifically drive the expression of the thymidine kinase (*tk*) gene in tumor cells was tested using a variety of constructs: One construct, designated pGRN266, contains an EcoRI-FseI PCR fragment with the *tk* gene cloned into the EcoRI-FseI sites of pGRN263. pGRN263, containing approximately 2.5 kb of hTERT promoter sequence, is similar to pGRN150, but contains a neomycin gene as selection marker. pGRN267 contains an EcoRI-FseI PCR fragment with the *tk* gene cloned into the EcoRI-FseI sites of pGRN264. pGRN264, containing approximately 210 bp of hTERT promoter sequence, is similar to pGRN176, but contains a neomycin gene as selection marker. pGRN268 contains an EcoRI-XbaI PCR fragment with the *tk* gene cloned into the EcoRI-XbaI (unmethylated) sites of pGRN265. pGRN265, containing approximately 90 bp of hTERT promoter sequence, is similar to pGRN175, but contains a neomycin gene as selection marker.

These hTERT promoter/*tk* constructs, pGRN266, pGRN267 and pGRN268, were re-introduced into mammalian cells and *tk*⁺ stable clones (and/or mass populations) were selected. Ganciclovir treatment in vitro of the *tk*⁺ cells resulted in selective destruction of all tumor lines tested, including 143B, 293, HT1080, Bxpc-3', DAOY and NIH3T3. Ganciclovir treatment had no effect on normal BJ cells.

Figure 5 is a map of the TPAC adenovector pGRN376. It was made by cloning the NOT1-BAMH1 fragment from pGRN267 into the NOT1-BGL2 sites of pAdBN (Quantum Biotech). The 7185 bp vector comprises the herpes simplex thymidine kinase (TK) gene under control of the medium-length hTERT promoter sequence.

Example 10: Transduction of hES cells with a thymidine kinase construct

These experiments test the effect of the pGRN376 vector described in the preceding Example on hES cells. The vector contains the herpes virus thymidine kinase gene under control of the telomerase reverse transcriptase promoter. Expression of the thymidine kinase gene in cells should render them susceptible to toxicity from the prodrug ganciclovir.

Undifferentiated H1 cells were plated into 24 well plates (1 confluent well of a 6 well plate split into 24 wells of a 24 well plate). After 48 h, some wells were infected with the TPAC vector at an MOI of 30 or 100. Four h after addition of the viral vector, medium was exchanged for new mouse embryonic fibroblast conditioned medium (mEF-CM); some wells received medium supplemented with 30 μ M ganciclovir (GCV). Cells exposed to GCV were re-fed with mEF-CM containing 30 μ M GCV daily for 4 days. On days 2,3, and 4 after the initiation of GCV treatment, wells were harvested and analyzed by flow cytometry to assess changes in 1) total cell number and 2) cell viability (measured by PI exclusion).

Figure 6 shows the results of this experiment. No change in total cell number was detected at MOI of 30 in the absence of GCV; but there was some decrease at MOI of 100 in absence of GCV starting at 48 h. Evidence for toxicity of GCV alone was detected: wells receiving GCV alone contained approximately 55% as many cells as the control wells on day 2, diminishing to 40% by day 4. Wells receiving GRN376 at MOIs of 30 or 100 cultured in the presence of GCV showed identical results: by day 2, these wells contained 18% of the cells contained in the control wells, while at days 3 and 4 these wells contained 6% and 8% of the cells in the control wells.

Slight toxicity was seen at MOI of 100 at day 4 in the absence of GCV (50% cells in ES gate vs. 83% for the control cells). Some toxicity of GCV alone was observed at d2, 75% cells in ES gate (vs. 85% control); at day 3, 68% (vs. 82% control); at day 4, 50% (vs. 65% control). Wells receiving GRN376 at MOIs of 30 or 100 cultured in the presence of GCV showed similar results: by day 2, these wells contained 24-28% cells in the ES gate, at day 3 they contained 19-22% cells in the ES gate, and at d4 these wells contained 12% cells in the ES gate. Thus, GRN376 plus GCV is effective at killing undifferentiated hES cells at an MOI as low as 30.

Titration Experiment

Undifferentiated H1 cells were plated into 24 well plates (1 confluent well of a 6 well plate split into 24 wells of a 24 well plate). After 48 h, some wells were infected with pGRN376 at an MOI of 30. Four h after addition of the viral vector, medium was exchanged for new mEF-CM; some wells received medium supplemented with 5, 10, 20, 30, or 40 μ M ganciclovir (GCV). Cells exposed to GCV were re-fed with mEF-CM containing GCV daily for 2 days. On day 2 after the initiation of GCV treatment, wells were harvested and analyzed by flow cytometry to assess changes in total cell number, and cell viability (measured by PI exclusion).

Figure 7 shows the results of this experiment. ~20 μ M GCV was optimal under the conditions tested.

Comparison of different hES lines

Undifferentiated hES of lines designated H1 and H7 cells were plated into 24 well plates (1 confluent well of a 6 well plate split into 24 wells of a 24 well plate). After 48 h, some wells were infected with pGRN376 at an MOI of 30. Four h after addition of the viral vector, medium was exchanged for new mEF-CM; some wells received medium supplemented with 20 μ M GCV. Cells exposed to GCV were re-fed with mEF-CM containing GCV daily for 3 days. On day 4 after the initiation of GCV treatment, wells were harvested and analyzed by flow cytometry to assess changes in total cell number.

Figure 8 shows the results. The total cell number demonstrated decreases in cell number for both lines after TPAC vector treatment. Thus, different hES cell lines respond to the TPAC vector. In subsequent studies, the H9 cell line was also found to be highly sensitive to GCV after TPAC vector treatment.

Example 11: Selection of differentiated cells

In this experiment, hES cells were treated with retinoic acid (RA) or dimethyl sulfoxide (DMSO), and then analyzed for hTERT and OCT-4 expression after treating with TPAC.

Undifferentiated H1 cells were plated into 24 well plates (1 confluent well of a 6 well plate split into 24 wells of a 24 well plate). 24 h later, some wells were re-fed with mEF-CM containing either 500 nM RA or 0.5% DMSO; wells were re-fed with medium supplemented with RA or DMSO for the remainder of the experiment. After 7 days of treatment with RA or DMSO, cells were infected with GRN376 at an MOI of 30.

Four h after addition of the viral vector, medium was exchanged for new mEF-CM (plus RA or DMSO where appropriate); some wells also received medium supplemented with 20 μ M ganciclovir (GCV). Cells exposed to GCV were re-fed with mEF-CM containing GCV daily for 3 days. On day 3 after the initiation of GCV treatment, wells were harvested and analyzed by flow cytometry to assess changes in total cell number. Additional wells were used in an effort to culture out any remaining undifferentiated stem cells; the medium of these wells was changed to mEF-CM (without RA, DMSO, or GCV). Cells were re-fed with mEF-CM every day for 7 days, then harvested for isolation of RNA. These samples were analyzed by quantitative RT-PCR for the expression of hTERT and OCT-4.

Figure 9 shows that the cell number decreased after TPAC treatment. After 7 days of drug pretreatment followed by TPAC plus GCV, all wells contained similar cell numbers. During the attempt to culture out surviving stem cells, the wells became confluent with highly differentiated appearing cells; no undifferentiated hES cells were obvious. Wells containing cells that had been pre-treated with RA were distinct in appearance from the cells either pre-treated with unadulterated mEF-CM or treated with mEF-CM plus DMSO. RT-PCR analysis (Lower Panel, non-quantitative, 35 cycles) showed that the surviving cells from the mEF-CM or DMSO-treated wells had no detectable OCT-4 expression, while 2 out of 4 RA-pre-treated samples presented very weak OCT-4 PCR products.

Thus, no detectable undifferentiated cells survive TPAC treatment followed by subsequent culture of wells grown in mEF-CM or mEF-CM plus DMSO. RA pre-treatment leads to detection of low levels of OCT-4 in the surviving cells. It is not clear whether this reflects persistence of undifferentiated stem cells or induction of another cell type that expresses OCT-4.

Example 12: Chimeric $\alpha(1,3)$ galactosyltransferase for optimal xenoantigen expression

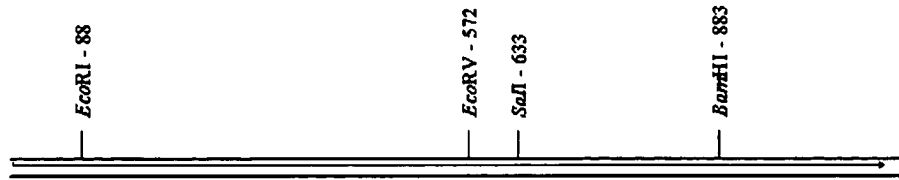
It is a hypothesis of this invention that elimination of undifferentiated cells will be enhanced and have greater efficiency with a higher density of xenodeterminant. When $\alpha(1,2)$ fucosyltransferase and $\alpha(1,3)$ galactosyltransferase are coexpressed, the $\alpha 1,2$ FT product dominates (US Patent 6166288). This is because the cytoplasmic domains of the two transferase enzymes direct them to different regions of the Golgi (Osman et al., supra). $\alpha 1,2$ FT is naturally expressed in virtually all metabolically active human cells.

To overcome this problem, a chimeric protein was designed in which a non-human $\alpha 1,3$ GT catalytic domain was joined to the human $\alpha 1,2$ FT cytoplasmic domain — thus producing a protein that should position itself in the Golgi in a more advantageous position.

The mouse $\alpha 1,3$ GT coding sequence was isolated from mouse kidney total RNA by RT/PCR using primers 5' of the start codon and 3' of the stop codon (5'- ggctgtac tacattgcctgga -3', SEQ. ID NO:14; 5'- gaaatagtgtca agttccatcaca -3', SEQ. ID NO:15). The PCR product was subcloned into pGEM T Easy™ (Promega) and used as a template for a PCR reaction to replace the cytoplasmic tail with a $\alpha 1,2$ FT cytoplasmic tail. The same 3' primer was used with an internal 5' primer containing the $\alpha 1,2$ FT codons (underlined) to replace the mouse $\alpha 1,3$ GT cytoplasmic tail (5'-cgatgtggctgag gagccaccggcag gtaatcctgttg atgctgattgtctc aac -3', SEQ. ID NO:16). The resulting PCR product was subcloned into pGEM T Easy™ (Promega).

MITMLQDLHVNKISMSRSKSETSLPSSRSGSQEKIMNVKGKVILLMLIVS	A (SEQ. ID NO:17)
<u>MWLRSHRQVILLMLIVS</u>	B (SEQ. ID NO:18)
<u>MWLRSHRQVVLSMLLV</u> S	C (SEQ. ID NO:19)
MNVKGRVVLSMLLV	D (SEQ. ID NO:20)

A: mouse $\alpha 1,3$ GT N terminal sequence
B: constructed chimeric mouse $\alpha 1,3$ GT N terminal sequence ($\alpha 1,2$ FT underlined)
C: chimeric porcine $\alpha 1,3$ GT N terminal sequence
D: porcine $\alpha 1,3$ GT N terminal sequence



**chimeric galT
1121 bp**

Example 13: $\alpha(1,3)$ galactosyltransferase under control of the endogenous hTERT promoter

Human ES cells are engineered to express $\alpha(1,3)$ galactosyltransferase ($\alpha 1,3$ GT) in several ways. The general strategy is illustrated in Figure 10.

The first option is a two-step targeting approach. A promoter trap targeting construct is generated to introduce a promoterless *neo* flanked by incompatible *lox* sites into intron 2 of the hTERT gene. The *lox-neo* cassette also contains a splice acceptor site adjacent to the 5'/*oxP* site. In this way, an mRNA is transcribed which encodes a fusion of neomycin phosphotransferase with the first few amino acids of the telomerase catalytic subunit to render targeted clones resistant to selection in G-418. The great majority of random integrants will not produce a transcript for G-418 resistance unless fortuitously integrated into an intron of an expressed gene. In a second step, the floxed *neo* is replaced with the ovine $\alpha 1,3$ GT encoding sequence by recombination mediated cassette exchange (RMCE) using a *cre* expressing plasmid (Lee et al., Gene 216 55, 1998; Kolb et al., Gene 227:21, 1999). Using the mutated *loxM* site eliminates recombination events in which *neo* is removed without replacement by $\alpha 1,3$ GT and ensures directionality of the insertion. Between the $\alpha 1,3$ GT and the 3'/*oxM* site is the tetracyclin-inducible promoter TetO7-CMVm and a fusion of the coding regions of exons 1 and 2 of the hTERT gene, such that second step targeting will (in the presence of transactivator) render the complete hTERT open reading frame inducible in tetracyclin or doxycycline (dox). In a subsequent step, targeted clones are transfected with the transactivator rtAVB, which binds the *tet* operator in TetO7-CMVm and induces transcription only in the presence of dox (Rossi et al. Nature Genetics 20 389, 1998).

In single step targeting experiments, *neo* is introduced with an internal ribosomal entry site (IRES) 3' to the $\alpha 1,3$ GT coding region. In this instance, the $\alpha 1,3$ GT encoding sequence is truncated before the polyadenylation signal and is transcribed directly from the hTERT promoter. The IRES allows transcription of a bicistronic message from which both proteins are translated. This single-step approach has the advantage that the initial modification is directly useful without a second cloning step. The two-step procedure, on the other hand, offers the advantage that a variety of open reading frames can be introduced to a single clone through *cre*-mediated recombination.

Three features of $\alpha 1,3$ GT-expressing hES cells are desirable: 1. That all undifferentiated hES cells are sensitive to human serum; 2. That all differentiated hES derivatives are resistant to human serum; 3. That in targeted clones, telomerization can be induced in vitro leading to increased proliferative life span.

Expression of $\alpha 1,3$ GT in modified hES cells is monitored by exposure of differentiated and undifferentiated cultures to fresh human serum (containing natural antibody to the Gal $\alpha(1,3)$ Gal epitope and

complement), and by RT-PCR. With appropriate levels of expression, there should be complement mediated lysis of undifferentiated cells, but no effect on differentiated cells. Based upon the frequency of random mutation at the HPRT locus in murine ES cells, a reversion frequency of approximately 10^{-7} can be expected. Hence, a similar background frequency among undifferentiated hES cells exposed to human serum indicates that cells containing the $\alpha 1,3$ GT transgene are effectively destroyed. Loss of $\alpha 1,3$ GT in surviving cells can be confirmed by Southern analysis. Thus, hES derived cells can be exposed to human serum in vitro immediately prior to transplantation, and selection against undifferentiated cells will continue to be active in vivo.

Clones in which the targeted hTERT allele is controlled by the TetO7-CMVm promoter are transfected with the transactivator rTAVB. Up to 10 rTAVB expressing clones derived from previously targeted subclones are assessed for responsiveness to induction and for telomere length. Expression and function of hTERT is monitored in vitro and compared with and without dox induction of the modified hTERT locus by RT-PCR and by TRAP assay, respectively.

References:

1. Insertion of a foreign gene into the β -casein locus by Cre-mediated site-specific recombination. Kolb, A., Ansell, R., McWhir J., Siddell, S. Gene Vol. 227:21-31, 1999.
2. Viable offspring derived from fetal and adult mammalian cells. Wilmut I, Schnieke AE, McWhir J, Kind AJ, Campbell KHS. Nature 385:10-813, 1997.
3. Selective ablation of differentiated cells permits isolation of embryonic stem cell lines from murine embryos with a non-permissive genetic background. McWhir J, Schieke AE, Ansell R, Wallace H, Colman A, Scott AR, Kind AJ. Nature Genetics, 14:223-226, 1996.
4. Sheep cloned by nuclear transfer from a cultured-cell line. Campbell KHS, McWhir J, Ritchie WA, Wilmut I. Nature, 380:64-66, 1996.
5. Use of double-replacement gene targeting to replace the murine alpha-lactalbumin gene with its human counterpart in embryonic stem-cells and mice. Stacey A, Schnieke A, McWhir J, Cooper J, Colman A, Melton DW. Mol. Cell Biol. 14:1009-1016, 1994.
6. Mice with DNA-repair gene (ercc-1) deficiency have elevated levels of p53, liver nuclear abnormalities and die before weaning. McWhir J, Selfridge J, Harrison DJ, Squires S, Melton DW. Nature Genetics 5:217-224, 1993.
7. In-vivo analysis of pim-1 deficiency. Laird PW, Vanderlugt NMT, Clarke A, Domen J, Linders K, McWhir J, Berns A, and Hooper M. Nucleic Acids Res. 21:4750-4755, 1993.
8. Gene targeting using a mouse hpvt minigene hpvt-deficient embryonic stem-cell system - inactivation of the mouse ercc-1 gene. Selfridge J, Pow AM, McWhir J, Magin TM, Melton DW. Somatic Cell Mol. Genet. 18:325-336, 1992.
9. Switching amino-terminal cytoplasmic domains of alpha(1,2)fucosyltransferase and alpha(1,3)galactosyltransferase alters the expression of H substance and Gal alpha(1,3)Gal. Osman, N. et al., J. Biol. Chem. 271:33105, 1996.

It will be recognized that the compositions and procedures provided in the description can be effectively modified by those skilled in the art without departing from the spirit of the invention embodied in the claims that follow.

SEQUENCE DATA

TABLE 2: Sequences listed in this Disclosure

SEQ. ID NO:	Designation	Reference
1	Lambda clone designated λ G ϕ 5 (ATCC Accession No. 98505) Contains human Telomerase Reverse Transcriptase (hTERT) genomic insert (residues 44-15375). The ATG translation initiation site begins at residue 13545.	GenBank Accession AF121948 International Patent Publication WO 00/46355.
2 - 5	Probes and Primers	(Artificial Sequences)
6	Sheep α 1,3-galactosyltransferase cDNA sequence	John Clark & Chris Denning; Geron Biomed
7	Sheep α 1,3-galactosyltransferase amino acid sequence	(<i>supra</i>)
8	Marmoset α 1,3-galactosyltransferase cDNA sequence	GenBank Accession S71333 See also Henion et al., Glycobiology 4,193 (1994)
9	Marmoset α 1,3-galactosyltransferase amino acid sequence	(<i>supra</i>)
10	Human histo blood group A transferase cDNA sequence	GenBank Accession J05175 See also Accession Nos. AF134413 & AF134412; Yamamoto et al., Nature 1990 May 17;345:229 (1990); U.S. Patent 5,326,857
11	Human histo blood group A transferase amino acid sequence	(<i>supra</i>)
12	Human histo blood group B transferase cDNA sequence	GenBank Accession AF134414 See also Yamamoto et al., Nature 1990 May 17;345:229 (1990); U.S. Patent 5,326,857
13	Human histo blood group B transferase amino acid sequence	(<i>supra</i>)
14-16	Probes and Primers	(Artificial Sequences)
17-20	Prototype α 1,3-galactosyltransferase catalytic domain/cytoplasmic domain fusions	This invention (Example 12)

SEQ. ID NO:1

```

GCGGCCCGCA GCTCTAATAC GACTCACTAT AGGGCGTCCA CTCGATCAAT GGAAGATGAG    60
GCATTGCCGA AGAAAAGATT AATGGATTTC AACACACAGC AACAGAAACT ACATGAAGTG    120
AAACACAGGA AAAAAAAGAT AAAGAAACGA AAAGAAAAGG GCATCAGTGA GCTTCAGCAG    180
AAGTTCCATC GGCCTTACAT ATGTGTAAGC AGAGGCCCTG TAGGAGCAGA GGCAGGGGGA    240
AAATACTTTA AGAAATAATG TCTAAAAGTT TTTCAAATAT GAGGAAAAAC ATAAAACAC    300
AGATCCAAGA AGCTCAACAA AACAAAGCAC AAGAAACAGG AAGAAATTAA AAGTTATATC    360
ACAGTCAAAT TGCTGAAAAC CAGCAACAAA GAGAATATCT TAAGAGTATC AGAGGAAAAG    420
AGATTAATGA CAGGCCAAGA AACAAATGAAA ACAATACAGA TTTCTTGTAG GAAACACAAG    480

```

ACAAAAGACA	TTTTTTAAAA	CCAAAAGGAA	AAAAAATGCT	ACATTAATAAT	GTTTTTTACC	540
CACTGAAAGT	ATATTTTCAA	ACATATTTTA	GGCCAGGCTT	GGTGGCTCAC	ACCTGTAATC	600
CCAGCACTTT	GGGAGGCCAA	GGTGGGTGGA	TGGCTTAAGG	TCAGGAGTTC	GAGACCAGCC	660
TGGCCAATAT	AGCGAAACCC	CATCTGTACT	AAAAACACAA	AAATTAGCTG	GGTGTGGTGA	720
CACATGCCTG	TAATCCCAAG	TACTCAGGAG	GCTAAGGCAG	GAGAAATTGCT	TGAACCTGGG	780
GGCAGAGGTG	GTGAGCCAAG	ATTGCACCAG	TGCACTCCAG	CCTTGGTGAC	AGAGTGAAAC	840
TCCATCTCAA	AAACAAACAA	ACAAAATACA	TATACATAAA	TATATATGCA	CATATATATA	900
CATATATAAA	TATATATACA	CATATATAAA	TCTATATACA	TATATACATA	TATACACATA	960
TATAAATCTA	TATACATATA	TATACATATA	TAATATATTT	ACATATATAA	ATATATACAT	1020
ATATAAATAT	ACATATATAA	ATACATATAT	AAATATACAT	ATATAAATAT	ACATATATAA	1080
ATATACATAT	ATAAATATAT	ACATATATAA	ATATACATAT	ATAAATATAT	ATACATATAT	1140
AAATATATAA	ATATACAAAT	ATATACAAAT	ATATACATAT	ATAAATGTAT	ATACGTATAT	1200
ACATATATAT	ATAAATATAT	AAAAAACTT	TTGGCTGGGC	ACCTTTCCAA	ATCTCATGGC	1260
ACATATAAGT	CTCATGGTAA	CCTCAAAATA	AAAAACATAT	AACAGATACA	CCAAAAATAA	1320
AAACCAATAA	ATTAATATCAT	GCCACCAGAA	GAAATTACCT	TCACTAAAAG	GAACACAGGA	1380
AGGAAAGAAA	GAAGGAAGAG	AAGACCATGA	AACAACCAGA	AAACAAACAA	CAAAACAGCA	1440
GGAGTAATTC	CTGACTTATC	AATAATAATG	CTGGGTGTAA	ATGGACTAAA	CTCTCCAATC	1500
AAAAGACATA	GAGTGGCTGA	ATGGACGAAA	AAAACAAGAC	TCAATAATCT	GTTGCCTACA	1560
AGAATATACT	TCACCTATAA	AGGGACACAT	AGACTGAAAA	TAAAAGGAAG	GAAAAATATT	1620
CTATGCAAAAT	GGAAACCAAA	AAAAGAACAG	AACTAGCTAC	ACTTATATCA	GACAAAAATAG	1680
ATTTCAAGAC	AAAAAGTACA	AAAAGAGACA	AAGTAATTAT	ATAATAATAA	AGCAAAAAAGA	1740
TATAACAAAT	TGGAATTTTAT	ATGGCCCAAA	CACCTGGGACA	CCCAGATATA	TACAGCAAAAT	1800
ATTATTAGAA	CTAAGGAGAG	AGAGAGATCC	CCATACAATA	ATAGCTGGAG	ACTTCACCCC	1860
GCTTTTAGCA	TTGGACAGAT	CATCCAGACA	GAAAATCAAC	CAAAAAATTG	GACTTAATCT	1920
ATAATATAGA	ACAAATGTAC	CTAATTGATG	TTTACAAGAC	ATTTTCATCCA	GTAGTTGCAG	1980
AATATGCATT	TTTTCTCAG	CATATGGATC	ATTCTCAAGG	ATAGACCATA	TATTAGGCCA	2040
CAGAACAAGC	CTATAAAAAAT	TCAAAAAAAT	TGAGCCAGGC	ATGATGGCTT	ATGCTTGTAA	2100
TTACAGCACT	TTGGGGAGGG	TGAGGTGGGA	GGATGTCTTG	AGTACAGGAG	TTTGAGACCA	2160
GCCTGGGCAA	AATAGTGAGA	CCCTGTCTCT	ACAAACTTTT	TTTTTTAATT	AGCCAGGCAT	2220
AGTGGTGTGT	GCCTGTAGTC	CCAGCTACTT	AGGAGGCTGA	AGTGGGAGGA	TCACTTGAGC	2280
CCAAGAGTTC	AAGGCTACGG	TGAGCCATGA	TTGCAACACC	ACACACCAGC	CTTGGTGACA	2340
GAATGAGACC	CTGTCTCAAA	AAAAAATAAA	AAAAATTGAAA	TAATATAAAG	CATCTTCTCT	2400
GGCCACAGTG	GAACAAAACC	AGAAATCAAC	AACAAGAGGA	ATTTTGAAAA	CTATACAAAC	2460
ACATGAAAAAT	TAAACAATAT	ACTTCTGAAT	AACCAGTGAG	TCAATGAAGA	AATTAATAAAG	2520
GAAATTGAAA	AATTTATTTA	AGCAAATGAT	AACGGAAACA	TAACTCTCA	AAACCCACGG	2580
TATACAGCAA	AAGCAGTGCT	AAGAAGGAAG	TTTATAGCTA	TAAGCAGCTA	CATCAAAAAA	2640
GTAGAAAAGC	CAGGCGCAGT	GGCTCATGCC	TGTAATCCCA	GCACTTTGGG	AGGCCAAGGC	2700
GGGCAGATCG	CCTGAGGTCA	GGAGTTCGAG	ACCAGCCTGA	CCAACACAGA	GAAACCTTGT	2760
CGCTACTAAA	AATACAAAAT	TAGCTGGGCA	TGGTGGCACA	TGCCTGTAAT	CCCAGCTACT	2820
CGGGAGGCTG	AGGCAGGATA	ACCGCTTGAA	CCCAGGAGGT	GGAGGTTGCG	GTGAGCCGGG	2880
ATTGCGCCAT	TGGACTCCAG	CCTGGGTAAC	AAGAGTGAAA	CCCTGTCTCA	AGAAAAAATA	2940
AAAAGTAGAA	AAACTTAAAA	ATACAACCTA	ATGATGCACC	TAAAGAACT	AGAAAAGCAA	3000
GAGCAAACCTA	AACCTAAAAT	TGGTAAAAGA	AAAGAAATAA	TAAAGATCAG	AGCAGAAATA	3060
AATGAAACTG	AAAGATAACA	ATACAAAAGA	TCAACAAAAT	TAAAAGTTGG	TTTTTTGAAA	3120
AGATAAAACA	AATTGACAAA	CCTTTGCCCA	GACTAAGAAA	AAAGGAAAGA	AGACCTAAAT	3180
AAATAAAGTC	AGAGATGAAA	AAAGAGACAT	TACAACCTGAT	ACCACAGAAA	TTCAAAGGAT	3240
CACTAGAGGC	TACTATGAGC	AACGTGACAC	TAATAAATTG	AAAAACCTAG	AAAAAATAGA	3300
TAAATTCCTA	GATGCATACA	ACCTACCAAG	ATTGAACCAT	GAAGAAATCC	AAAGCCCAAA	3360
CAGACCAATA	ACAATAATGG	GATTAAGGCC	ATAATAAAAA	GTCTCCTAGC	AAAGAGAAGC	3420
CCAGGACCCA	ATGGCTTCCC	TGCTGGATT	TACCAATCAT	TAAAGAAAGA	ATGAATTCCA	3480
ATCCTACTCA	AACTATTCTG	AAAAATAGAG	GAAAGAATAC	TTCCAAACTC	ATTCTACATG	3540
GCCAGTATTA	CCCTGATTCC	AAAACAGAC	AAAAACACAT	CAAAAAACAA	CAAAACAAAA	3600
AACAGAAAGA	AAGAAAACCTA	CAGGCCAATA	TCCCTGATGA	ATACTGATAC	AAAAATCCTC	3660
AACAAAACAC	TAGCAAAACA	AATTAACAAA	CACCTTCGAA	AGATCATTCA	TTGTGATCAA	3720
GTGGGATTTA	TTCCAGGGAT	GGAAGGATGG	TTCAACATAT	GCAAAATCAAT	CAATGTGATA	3780
CATCATCCCA	ACAAAATGAA	GTACAAAAAC	TATATGATTA	TTTCACTTTA	TGCAGAAAAA	3840
GCATTTTGATA	AAATTCTGCA	CCCTTCATGA	TAAAAACCTT	CAAAAAACCA	GGTATACAAG	3900
AAACATACAG	GCCAGGCACA	GTGGCTCACA	CCTGCGATCC	CAGCACTCTG	GGAGGCCAAG	3960
GTGGGATGAT	TGCTTGGGCC	CAGGAGTTTG	AGACTAGCCT	GGGCAACAAA	ATGAGACCTG	4020
GTCTACAAAA	AACTTTTTTTA	AAAAATTAGC	CAGGCATGAT	GGCATATGCC	TGTAGTCCCA	4080
GCTAGTCTGG	AGGCTGAGGT	GGGAGAATCA	CTTAAGCCTA	GGAGGTCGAG	GCTGCAGTGA	4140
GCCATGAACA	TGTCACTGTA	CTCCAGCCTA	GACAACAGAA	CAAGACCCCA	CTGAATAAGA	4200
AGAAGGAGAA	GGAGAAGGGA	GGAAGGAGGG	AGAAGGGAGG	AGGAGGAGAA	GGAGGAGGTG	4260
GAGGAGAAGT	GGAAGGGGAA	GGGGAAGGGA	AAGAGGAAGA	AGAAGAAACA	TATTTTCAACA	4320
TAATAAAAGC	CCTATATGAC	AGACCGAGGT	AGTATTATGA	GGAAAAACTG	AAAGCCTTTC	4380
CTCTAAGATC	TGGAAAATGA	CAAGGCCCA	CTTTCACCAC	TGTGATTCAA	CATAGTACTA	4440
GAAGTCCTAG	CTAGAGCAAT	CAGATAAGAG	AAAGAAATAA	AAGGCATCCA	AACTGGAAAG	4500
GAAGAAGTCA	AATTATCCTG	TTTGACAGATG	ATATGATCTT	ATATCTGGAA	AAGACTTAAG	4560

ACACCACTAA	AAAACATTA	GAGCTGAAAT	TTGGTACAGC	AGGATACAAA	ATCAATGTAC	4620
AAAAATCAGT	AGTATTTCTA	TATTCCAACA	GCAAAACAATC	TGAAAAAGAA	ACCAAAAAAG	4680
CAGCTACAAA	TAAAAATTA	CAGCTAGGAA	TTAACCAAG	AAGTGAAAGA	TCTCTACAA	4740
GAAAACTATA	AAATATTGAT	AAAAGAAAT	GAAGAGGGCA	CAAAAAAGA	AAAGATATTC	4800
CATGTTTATA	GATTGGAAGA	ATAAATACTG	TTAAAAATGTC	CATACTACCC	AAAGCAATTT	4860
ACAAATTC	TGCAATCCCT	ATTAATAATAC	TAATGACGTT	CTTCACAGAA	ATAGAAGAAA	4920
CAATTC	ATTTGTACAG	AACCACAAAA	GACCCAGAAT	AGCCAAAGCT	ATCCTGACCA	4980
AAAAGAACA	AACTGGAAGC	ATCACATTAC	CTGACTTCAA	ATTATACTAC	AAAGCTATAG	5040
TAACCAAC	TACATGGTAC	TGGCATAAAA	ACAGATGAGA	CATGGACCAG	AGGAACAGAA	5100
TAGAGAAATCC	AGAAACAAAT	CCATGCATCT	ACAGTGAAT	CATTTTGTAC	AAAGGTGCCA	5160
AGAACATACT	TTGGGGAAAA	GATAATCTCT	TCAATAAATG	GTGCTGGAGG	AACTGGATAT	5220
CCATATGCAA	AATAACAATA	CTGAACCTCT	GTCTCTCACC	ATATACAAAA	GCAAAATCAAA	5280
ATGGATGAAA	GGCTTAAATC	TAAAACCTCA	AACTTTGCAA	CTACTAAAAG	AAAACACCGG	5340
AGAAACTCTC	CAGGACATTG	GAGTGGGCAA	AGACTTCTTG	AGTAATTTCC	TGCAGGCACA	5400
GGCAACCAAA	GCAAAAACAG	ACAAATGGGA	TCAATATCAAG	TTAAAAAGCT	TCTGCCCAGC	5460
AAAGGAAACA	ATCAACAAAG	AGAAGAGACA	ACCCACAGAA	TGGGAGAATA	TATTTGCAAA	5520
CTATTCACTCT	AACAAGGAAT	TAATAACCAG	TATATATAAG	GAGCTCAAAC	TACTCTATAA	5580
GAAAAACACC	TAATAAGCTG	ATTTTCAAAA	ATAAGCAAAA	GATCTGGGTA	GACATTTCTC	5640
AAAATAAGTC	ATACAAATGG	CAACACGGCA	TCTGAAAATG	TGCTCAACAC	CACTGATCAT	5700
CAGAGAAATG	CAAAATCAAAA	CTACTATGAG	AGATCATCTC	ACCCACAGTTA	AAATGGCTTT	5760
TATTTCAAAAG	ACAGGCAATA	ACAAATGCCA	GTGAGGATGT	GGATAAAAGG	AAACCTTTGG	5820
ACACTGTTGG	TGGGAATGGA	AATTGCTACC	ACTATGGAGA	ACAGTTTGAA	AGTTCTCTCA	5880
AAAACTAAAA	ATAAAGCTAC	CATACAGCAA	TCCCATTTGCT	AGGTATATAC	TCCAAAAAAG	5940
GGAATCAGTG	TATCAACAAG	CTATCTCCAC	TCCCACATTT	ACTGCAGCAC	TGTTCTATAGC	6000
AGCCAAGGTT	TGGAAGCAAC	CTCAGTGTC	ATCAACAGAC	GAATGGAAAA	AGAAAATGTG	6060
GTGCACATAC	ACAATGGAGT	ACTACGCAGC	CATAAAAAAG	AATGAGATCC	TGTCAGTTGC	6120
AACAGCATGG	GGGGCACTGG	TCAGTATGTT	AAGTGAAATA	AGCCAGGCAC	AGAAAGACAA	6180
ACTTTTCATG	TTCTCCCTTA	CTTGTGGGAG	CAAAAAATTAA	AACAATTGAC	ATAGAAATAG	6240
AGGAGAAATGG	TGGTCTAGA	GGGGTGGGGG	ACAGGGTGAC	TAGAGTCAAC	AATAATTTAT	6300
TGTATGTTTT	AAAATAACTA	AAAGAGTATA	ATTGGGTTGT	TTGTAACACA	AAGAAAGGAT	6360
AAATGCTTGA	AGGTGACAGA	TACCCCATTT	ACCCTGATGT	GATTATTACA	CATTGTATGC	6420
CTGTATCAAA	ATATCTCATG	TATGCTATAG	ATATAAACCC	TACTATATTA	AAAAATTAATA	6480
TTTTAATGGC	CAGGCACGGT	GGCTCATGTC	CATAATCCCA	GCACTTTGGG	AGGCCGAGGC	6540
GGTGGATCAC	CTGAGGTGAG	GAGTTTGAAA	CCAGTCTGGC	CACCATGATG	AAACCTGTCT	6600
TCTACTAAAG	ATACAAAAAT	TAGCCAGGCG	TGGTGGCACA	TACCTGTAGT	CCCACTACT	6660
CAGGAGGCTG	AGACAGGAGA	ATTGCTTGAA	CCTGGGAGGC	GGAGGTTGCA	GTGAGCCGAG	6720
ATCATGCCAC	TGCACTGCAG	CCTGGGTGAC	AGAGCAAGAC	TCCATCTCAA	AACAAAAACA	6780
AAAAAAAGAA	GATTAATAAT	GTAATTTTTA	TGTACCGTAT	AAATATATAC	TCTACTATAT	6840
TAGAAGTTAA	AAATTAATAAC	AATTATAAAA	GGTAATTAAC	CACCTAATCT	AAAAATAAGAA	6900
CAATGTATGT	GGGGTTTCTA	CGTTCTGAAG	AAGTAAAAAGT	TATGGCCACG	ATGGCAGAAA	6960
TGTGAGGAGG	GAACAGTGGG	AGTTACTGTT	GTTAGACGCT	CATACTCTCT	GTAAGTGACT	7020
TAATTTTAAC	CAAAGACAGG	CTGGGAGAAG	TTAAAGAGGC	ATTCTATAAG	CCCTAAAAACA	7080
ACTGCTAATA	ATGGTGAAAG	GTAATCTCTA	TTAATTACCA	ATAATTACAG	ATATCTCTAA	7140
AATCGAGCTG	CAGAATTGGC	ACGTCTGATC	ACACCGTCT	CTCATTACAG	GTGCTTTTTT	7200
TCTTGTGTGC	TTGGAGATTT	TCGATTGTGT	GTTCTGTGTT	GGTTAACTT	AATCTGTATG	7260
AATCCTGAAA	CGAAAAATGG	TGGTGATTTT	CTCCAGAAGA	ATTAGAGTAC	CTGGCAGGAA	7320
GCAGGTGGCT	CTGTGGACCT	GAGCCACTTC	AATCTTCAAG	GGTCTCTGGC	CAAGACCCAG	7380
GTGCAAGGCA	GAGGCCTGAT	GACCCGAGGA	CAGGAAAGCT	CGGATGGGAA	GGGGCGATGA	7440
GAAGCCTGCC	TCGTTGGTGA	GCAGCGCATG	AAGTGCCCTT	ATTTACGCTT	TGCAAGATT	7500
GCTCTGATA	CCATCTGGAA	AAGGCGGCCA	GCGGGAATGC	AAGGAGTCAG	AAGCCTCTG	7560
CTCAAACCCA	GGCCAGCAGC	TATGGCGCCC	ACCCGGGCGT	GTGCCAGAGG	GAGAGGAGTC	7620
AAGGCACCTC	GAAGTATGGC	TTAAATCTTT	TTTTACCTG	AAGCAGTGAC	CAAGGTGTAT	7680
TCTGAGGGAA	GCTTGAGTTA	GGTGCCTTCT	TTAAACAGA	AAGTCATGGA	AGCACCTTCT	7740
TCAAGGGAAA	ACCAGACGCC	CGCTCTGCGG	TCATTTACCT	CTTCTCTCTC	TCCCTCTCTT	7800
GCCCTCGCGG	TTTCTGATCG	GGACAGAGTG	ACCCCGTGG	AGCTTCTCCG	AGCCCGTGCT	7860
GAGGACCCTC	TTGCAAAAGG	CTCCACAGAC	CCCCGCCCTG	GAGAGAGGAG	TCTGAGCCTG	7920
GCTTAATAAC	AAACTGGGAT	GTGGCTGGGG	GCGGACAGCG	ACGGCGGGAT	TCAAAGACTT	7980
AATTCCATGA	GTAATTCAA	CCTTTCCACA	TCCGAATGGA	TTTGGATTTT	ATCTTAATAT	8040
TTTCTTAAT	TTATCAAAAT	AACATTCAAG	AGTGACAGAAA	TCCAAAGCGG	TAAAAACAGGA	8100
ACTGAGCTAT	GTTTGCCAAG	GTCCAAGGAC	TTAATAACCA	TGTTCAAGAG	GATTTTTTCGC	8160
CCTAAGTACT	TTTTATTGGT	TTTCATAAGG	TGGCTTAGGG	TGCAAGGGAA	AGTACACGAG	8220
GAGAGGACTG	GGCGGAGGG	CTATGAGCAC	GGCAAGGCCA	CCGGGAGAG	AGTCCCCGGC	8280
CTGGGAGGCT	GACAGCAGGA	CCACTGACCG	TCCTCCCTGG	GAGCTGCCAC	ATTGGGCAAC	8340
GCGAAGGCGG	CCAGCTGGC	TGTGACTCAG	GACCCCATAC	CGGCTTCTG	GGCCCAACCA	8400
CACTAACCCA	GGAAGTCACG	GAGCTCTGAA	CCCGTGGA	CGAACATGAC	CCTTGCCTGC	8460
CTGCTTCCCT	GGGTGGGTCA	AGGGTAATGA	AGTGGTGTGC	AGGAAATGGC	CATGTAAATT	8520
ACACGACTCT	GCTGATGGGG	ACCGTTCCTT	CCATCATTAT	TCATCTTCAC	CCCCAAGGAC	8580
TGAATGATTC	CAGCAACTTC	TTCGGGTGTG	ACAAGCCATG	ACAACACTCA	GTACAAACAC	8640

CACTCTTTTA	CTAGGCCAC	AGAGCACGGC	CCACACCCCT	GATATATTAA	GAGTCCAGGA	8700
GAGATGAGGC	TGCTTTTCCAG	CACCAGGCTG	GGGTGACAAC	AGCGGCTGAA	CAGTCTGTTC	8760
CTCTAGACTA	GTAGACCCCTG	GCAGGCACTC	CCCCAGATTG	TAGGGCCTGG	TTGCTGCTTC	8820
CCGAGGGGCG	CATCTGCCCT	GGAGACTCAG	CCTGGGGTGC	CACACTGAGG	CCAGCCCTGT	8880
CTCCACACCC	TCCGCCTCCA	GGCCTCAGCT	TCTCCAGCAG	CTTCCTAAAC	CCTGGGTGGG	8940
CCGTGTTCCA	GCGCTACTGT	CTCACCTGTC	CCACTGTGTC	TTGTCTCAGC	GACGTAGCTC	9000
GCACGGTTCC	TCCTCACATG	GGGTGTCTGT	CTCCTTCCCC	AACACTCACA	TGCGTTGAAG	9060
GGAGGAGATT	CTGCGCCTCC	CAGACTGGCT	CCTCTGAGCC	TGAACCTGGC	TCGTGGCCCC	9120
CGATGCAGGT	TCCTGGCGTC	CGGCTGCACG	CTGACCTCCA	TTTCCAGGCG	CTCCCCGTCT	9180
CCTGTCTAT	GCCGGGGCCT	GCCGGTGTGT	TCCTCTGTTT	CTGTGCTCCT	TTCCACGTCC	9240
AGCTGCGTGT	GTCTCTGTCC	GCTAGGGTCT	CGGGGTTTTT	ATAGGCATAG	GACGGGGGCG	9300
TGGTGGGCCA	GGGCGCTCTT	GGGAAATGCA	ACATTTGGGT	GTGAAAGTAG	GAGTGCCTGT	9360
CCTCACCTAG	GTCCACGGGC	ACAGGCTGCG	GGATGGAGCC	CCCCCAGGG	ACCCGCCCTT	9420
CTCTGCCAG	CACCTTTCTG	CCCCCTCCC	TCTGGAACAC	AGAGTGGCAG	TTTCCACAAG	9480
CACTAAGCAT	CCTCTTCCCA	AAAGACCCAG	CATTGGCACC	CCTGGACATT	TGCCCCACAG	9540
CCCTGGGAAT	TCACGTGACT	ACGCACATCA	TGTACACACT	CCCGTCCACG	ACCGACCCCC	9600
GCTGTTTTAT	TTTAAATAGCT	ACAAAGCAGG	GAAATCCCTG	CTAAAATGTC	CTTTAAACAAA	9660
CTGGTTAAAC	AAACGGGTCC	ATCCGCACGG	TGGACAGTTC	CTCACAGTGA	AGAGGAACAT	9720
GCCGTTTATA	AAGCCTGCAG	GCATCTCAAG	GGAAATTACG	TGAGTCAAAA	CTGCCACCTC	9780
CATGGGATAC	GTACGCAACA	TGCTCAAAAA	GAAAGAATTT	CACCCCATGG	CAGGGGAGTG	9840
GTTGGGGGGT	TAAGGACGGT	GGGGGCAGCA	GCTGGGGGCT	ACTGCACGCA	CCTTTTACTA	9900
AAGCCAGTTT	CTTGGTTCTG	ATGGTATTGG	CTCAGTTATG	GGAGACTAAC	CATAGGGGAG	9960
TGGGGATGGG	GGAACCCGGA	GGCTGTGCCA	TCTTTGCCAT	GCCCGAGTGT	CCTGGGCAGG	10020
ATAATGCTCT	AGAGATGCCC	ACGTCTGAT	TCCCCAAAC	CTGTGGACAG	AACCCGCCCCG	10080
GCCCCAGGGC	CTTTGCAGGT	GTGATCTCCG	TGAGGACCTT	GAGGTCTGGG	ATCCTTCGGG	10140
ACTACCTGCA	GGCCGAAAA	GTAATCCAGG	GGTTCTGGGA	AGAGGCGGGC	AGGAGGGTCA	10200
GAGGGGGGCA	GCCTCAGGAC	GATGGAGGCA	GTCAGTCTGA	GGCTGAAAAG	GGAGGGAGGG	10260
CCTCGAGCCC	AGGCCTGCAA	GCGCTCCAG	AAGCTGGAAG	AAGCGGGGAA	GGGACCTCTC	10320
ACGGAGCCTG	CAGCAGGAAG	GCACGCTGG	CCCTTAGCCC	ACCAGGGCCC	ATCGTGGACC	10380
TCCGGCCTCC	GTGCCATAGG	AGGGCACTCG	CGCTGCCCTT	CTAGCATGAA	GTGTGTGGGG	10440
ATTTGCAGAA	GCAACAGGAA	ACCCATGCAC	TGTGAATCTA	GGATTATTTT	AAAACAAAGG	10500
TTTACAGAAA	CATCCAAGGA	CAGGGCTGAA	GTGCCTCCGG	GCAAGGGCAG	GGCAGGCACG	10560
AGTGATTTTA	TTTAGCTATT	TTATTTTATT	TACTTACTTT	CTGAGACAGA	GTTATGCTCT	10620
TGTTGCCAG	GCTGGAGTGC	AGCGGCATGA	TCTTGGCTCA	CTGCAACCTC	CGTCTCCTGG	10680
GTTCAAGCAA	TTCTCGTGCC	TCAGCCTCCC	AAGTAGCTGG	GATTTTCAGG	GTGCACCACC	10740
ACACCCGGCT	AATTTGTAT	TTTTAGTAGA	GATGGGCTTT	CACCATGTTG	GTGAGGCTGA	10800
TCTCAAAATC	CTGACCTCAG	GTGATCCGCC	CACCTCAGCC	TCCCAAAGTG	CTGGGATTAC	10860
AGGCATGAGC	CACCTGCACCT	GGCCTATTTA	ACCATTTTAA	AACCTTCCCTG	GGCTCAAGTC	10920
ACACCCACTG	GTAAGGAGTT	CATGGAGTTC	AATTTCCCTT	TACTCAGGA	GTTACCCTCC	10980
TTTGATATT	TCTGTAATTC	TTCTGATAGC	GGGATACAC	CGTCTCTTGA	CATATTACCA	11040
GTTTCTGTGA	CCACCTGTTA	TCCCATGGGA	CCCACTGCAG	GGGCAGCTGG	GAGGCTGCAG	11100
GCTTCAGGTC	CCAGTGGGGT	TGCCATCTGC	CAGTAGAAAC	CTGATGTAGA	ATCAGGGCGC	11160
GAGTGTGGAC	ACTGTCCTGA	ATCTCAATGT	CTCAGTGTGT	GCTGAAACAT	GTAGAAATTA	11220
AAGTCCATCC	CTCCTACTCT	ACTGGGATTG	AGCCCTTCC	CTATCCCCC	CCAGGGGCGG	11280
AGGAGTTTCT	CTCCTCCTG	TGGAGCAAGG	AATGATACTT	TGTTATTTT	CACTGCTGGT	11340
ACTGAATCCA	CTGTTTCATT	TGTTGGTTTG	TTTGTTTTGT	TTTGAGAGGC	GGTTTCACTC	11400
TTGTTGTCTA	GGCTGGAGGG	AGTGCAATGG	CGCGATCTTG	GCTTACTGCA	GCCTCTGCCT	11460
CCCAGGTTCA	AGTGATTCTC	CTGCTTCCGC	CTCCATTG	GCTGGGATTA	CAGGCACCCG	11520
CCACCATGCC	CAGCTAATTT	TTTGTATTTT	TAGTAGAGAC	GGGGGTGGGG	GTGGGGTTCA	11580
CCATGTTGGC	CAGGCTGGTC	TGCAACTTCT	GACCTCAGAT	GATCCACCTG	CCTCTGCCTC	11640
CTAAAGTGCT	GGGATTACAG	GTGTGAGCCA	CCATGCCCAG	CTCAGAATTT	ACTCTGTTTA	11700
GAAACATCTG	GGTCTGAGGT	AGGAAAGTCA	CCCCACTCAA	GTGTTGTGGT	GTTTTAAGCC	11760
AATGATAGAA	TTTTTTTATT	GTTGTTAGAA	CACCTTGTAT	GTTTTACACT	GTGATGACTA	11820
AGACATCATC	AGCTTTTCAA	AGACACACTA	ACTGCACCCA	TAATACTGGG	GTGCTTCTG	11880
GGTATCAGCG	ATCTTCATTG	AATGCCGGGA	GGCGTTTCT	CGCCATGCAC	ATGGGTGTTAA	11940
TTACTCCAGC	ATAATCTTCT	GCTTCCATTT	CTTCTCTTCC	CTCTTTTAAA	ATTGTGTTTT	12000
CTATGTTGGC	TTCTCTGCAG	AGAACCAGTG	TAAGCTACAA	CTTAACTTTT	GTTGGAACAA	12060
ATTTTCCAAA	CCGCCCCCTT	GCCCTAGTGG	CAGAGACAA	TCACAAACAC	AGCCCTTTAA	12120
AAAGGCTTAG	GAGTCACTAA	GGGATTTTCT	AGAAGAGCGA	CCCGTAATCC	TAAGTATTTA	12180
CAAGACGAGG	CTAACCTCCA	GCGAGCGTGA	CAGCCCAGGG	AGGGTGGCAG	GCCTGTTCAA	12240
ATGCTAGCTC	CATAAATAAA	GCAATTTCT	CCGGCAGTTT	CTGAAAGTAG	GAAAGGTTAC	12300
ATTTAAGGTT	GCGTTTGTTA	GCAATTCAGT	GTTTGCCGAC	CTCAGCTACA	GCATCCCTGC	12360
AAGGCCTCGG	GAGACCCAGA	AGTTTCTCGC	CCCTTAGATC	CAAACTTGAG	CAACCCGGAG	12420
TCTGGATTCC	TGGGAAGTCC	TCAGCTGTCC	TGCGGTTGTG	CCGGGGCCCC	AGGTCTGGAG	12480
GGGACCACTG	GCCGTGTGGC	TTCTACTGCT	GGGCTGGAAG	TCGGGCTCC	TAGCTCTGCA	12540
GTCCGAGGCT	TGGAGCCAGG	TGCCTGGACC	CCGAGGCTGC	CCTCCACCCT	GTGCGGGCGG	12600
GATGTGACCA	GATGTTGGCC	TCATCTGCCA	GACAGAGTGC	CGGGGCCACG	GCTCAAGGCC	12660
GTTGTGGCTG	GTGTGAGGCG	CCCCGTGGCG	GGCCAGCAGG	AGCGCCTGGC	TCCATTTCCC	12720

ACCCTTTCTC	GACGGGACCG	CCCCGGTGGG	TGATTAACAG	ATTTGGGGTG	GTTTGCTCAT	12780
GGTGGGGACC	CCTCGCCGCC	TGAGAACCTG	CAAAGAGAAA	TGACGGGCCT	GTGTCAAGGA	12840
GCCCAAGTCG	CGGGGAAGTG	TTGCAGGGAG	GCACTCCGGG	AGGTCCCGCG	TGCCCCGTCCA	12900
GGGAGCAATG	CGTCTCTGGG	TTCTGTCCTA	GCCGCGTCTA	CGCGCTCCG	TCCTCCCTTT	12960
CACGTCCGGC	ATTCGTGGTG	CCCGAGCCCG	GACGCCCCCG	GTCCGGACCT	GGAGGCAGCC	13020
CTGGGTCTCC	GGATCAGGCC	AGCGGCCAAA	GGGTGCGCCG	ACGCACCTGT	TCCCAGGGCC	13080
TCCACATCAT	GGCCCCCTCC	TCCGGTTACC	CCACAGCCTA	GGCCGATTCT	ACCTCTCTCC	13140
GCTGGGGCCC	TCGCTGGCGT	CCCTGCACCC	TGGGAGCGCG	AGCGGCGCGC	GGGCGGGGAA	13200
GCGCGGCCCA	GACCCCGGGG	TCCGCGCGGA	GCAGCTGCGC	TGTGCGGGCC	AGGCCGGGCT	13260
CCCAGTGGAT	TCGCGGGCAC	AGACGCCCCG	GACCGCGCTT	CCCACGTGGC	GGAGGGACTG	13320
GGGACCCGGG	CACCCGTCCT	GCCCTTTCAC	CTTCCAGCTC	CGCCTCTCTC	GCGCGGACCC	13380
CGCCCCGTCC	CGACCCCTCC	CGGGTCCCCG	GCCACGCCCC	CTCCGGGCCC	TCCCAGCCCC	13440
TCCCTTTCTC	TTCCGCGGCC	CGCCCTCTCT	CTCGCGGCGC	GAGTTTTCAGG	CAGCGCTGCG	13500
TCCTGCTGCG	CACGTGGGAA	GCCCTGGCCC	CGGCCACCCC	CGCGATGCGG	CGCGTCCCCC	13560
GCTGCCGAGC	CGTGCGCTCC	CTGCTGCGCA	GCCACTACCG	CGAGGTGCTG	CCGCTGGCCA	13620
CGTTCGTGCG	GCGCCTGGGG	CCCCAGGGGT	GGCGGTGGT	GCAGCGCGGG	GACCCGGCGG	13680
CTTTCGCGCG	GCTGCTGGCC	CAGTGCCCTG	TGTGCGTGCC	CTGGGACGCA	CGGCCGCCCC	13740
CCGCGGCCCC	CTCCTTCCGC	CAGGTGGGGC	TCCCCGGGGT	CGGCGTCCGG	CTGGGGTTGA	13800
GGGCGGCGCG	GGGGAACCAG	CGACATGCGG	AGAGCAGCGC	AGGCGACTCA	GGGCGCTTCC	13860
CCCGCAGGTG	TCCTGCTCTA	AGGAGCTGGT	GGCCCGAGTG	CTGCAGAGGC	TGTGCGAGCG	13920
CGGCGCGAAG	AACGTGCTGG	CCTTCGGCTT	CGCGCTGCTG	GACGGGGCCC	GCGGGGGCCC	13980
CCCCAGGGCC	TTACACCA	CGGTGCGCAG	CTACCTGCCC	AACACGGTGA	CCGACGCACT	14040
GCGGGGAGCG	GGGGCGTGGG	GGCTGCTGCT	GCGCCGCGTG	GGCGACGACG	TGCTGGTTCA	14100
CCTGCTGGCA	CGCTGCGCGC	TCTTTGTGCT	GGTGGTCCC	AGCTGCGCCT	ACCAGGTGTG	14160
CGGGCGCGCG	CTGTACCAGC	TGCGCGCTGC	CACTCAGGCC	CGGCCCCCGC	CACACGCTAG	14220
TGGACCCCGA	AGGCGTCTGG	GATCGGAACG	GGCTGGAAC	CATAGCGTCA	GGGAGGCCGG	14280
GGTCCCCCTG	GGCCTGCCAG	CCCCGGGTGC	GAGGAGGCGC	GGGGGCACTG	CCAGCCGAAG	14340
TCTGCCGTTG	CCCAAGAGGC	CCAGGCGTGG	CGCTGCCCT	GAGCCGGAGC	GGAGCCCGCT	14400
TGGGACAGGG	TCCTGGGCCC	ACCCGGGCGG	GACGCGTGGA	CCGAGTGACC	GTGGTTTCTG	14460
TGTGGTGTC	CCTGCCAGAC	CCGCCGAAGA	AGCCACCTCT	TTGGAGGGTG	CGCTCTCTGG	14520
CACGCCGCAC	TCCCACCCAT	CCGTGGGCGG	CCAGCACAC	GCGGGCCCCC	CATCCACATC	14580
GCGGCCACCA	CGTCCCTGGG	ACACGCCCTT	TCCCCCGGTG	TACGCCGAGA	CCAAGCACTT	14640
CCTCTACTCC	TCAGGCGACA	AGGAGCAGCT	GCGGCCCTCC	TTCTACTCA	GCTCTCTGAG	14700
GCCCAGCCTG	ACTGGCGCTC	GGAGGCTCGT	GGAGACCATC	TTTCTGGGTT	CCAGGCCCTG	14760
GATGCCAGGG	ACTCCCCGCA	GGTTGCCCGG	CCTGCCCCAG	CGCTACTGGC	AAATGCCGGC	14820
CCTGTTTCTG	GAGCTGCTTG	GGAACCAAGC	GCAGTGCCCC	TACGGGGTGC	TCCTCAAGAC	14880
GCACTGCCCC	CTGCGAGCTG	CGGTACCCCC	AGCAGCCGGT	GTCTGTGCCC	GGGAGAAGCC	14940
CCAGGGCTCT	GTGGCGGCCC	CCGAGGAGGA	GGACACAGAC	CCCCGTGCCC	TGGTGACGCT	15000
GCTCCGCCAG	CACAGCAGCC	CCTGGCAGGT	GTACGGCTTC	GTGCGGGCCT	GCCTGCGCCG	15060
GCTGGTGCCC	CCAGGCTCT	GGGGTCCAG	GCACAACGAA	CGCCGCTTCC	TCAGGAACAC	15120
CAAGAAGTTC	ATCTCCCTGG	GGAAGCATGC	CAAGCTCTCG	CTGCAGGAGC	TGACGTGGAA	15180
GATGAGCGTG	CGGGACTGCG	CTTGGCTGCG	CAGGAGCCCA	GGTGAGGAGG	TGGTGCCCGT	15240
CGAGGGCCCA	GGCCCCAGAG	CTGAATGCAG	TAGGGGCTCA	GAAAAGGGGG	CAGGCAGAGC	15300
CCTGGTCCTC	CTGTCTCCAT	CGTCACGTGG	GCACACGTGG	CTTTTCGCTC	AGGACGTCGA	15360
GTGGACACGG	TGATCGAGTC	GACTCCCTTT	AGTGAGGGTT	AATTGAGCTC	GCGGCCCC	15418

SEQ. ID NO:6

AGCCGAGGACGCCCGCGGGGAGCCGAGGCTCCGGCCAGCCCCAGCGCGCCAGCTTCTG
CAGATCAGG
AGTCAGAACGCTGCAC
CTTCGCTTCTCCAGCCCTGCCTCCTTCTGCAAAACGGAGCTCAATAGAACTTGGTACT
TTTGCCCTTTTACTCTGGGAGGAGAGAAGCAGCATGAGGAGAAAAATA
[beginning of codins sequence]
ATGAATGTCAAA
GGAAAAGTGATTCTGTCAATGCTGGTTGTCTCAACTGTCAATTGTTGTGTTTGGGAATAT
ATCCACAGCCAGAGGCTCTTTGTTCTGGATAAACCCATCAAGAAACCCAGAAAGTCAGT
GGCGGCAGCAGCATTAGAGGGCTGGTGGTTTCCGAGATGGTTTAACAATGGTTACCAA
GAAGAAGATGAAGACGTAGACGAAGAAAAGGAACAAAGAAAGGAAGACAAAGCAAGCTT
AAGCTATCGGACTGGTTCAACCCATTTAAACGCCCTGAGGTTGTGACTATGACAGATTGG
AAGGCACCCGCTGGTGTGGGAAGGCACTTACACAGAGCCGCTTAGACGATTACTACGCC
AAGCAGAAAAATTACCGTCGGCTGACGGTTTTTCGCCGTCGGAAGATACATTGAGCATTAC
TTGGAGGATTTCTTAACGTCTGCTAATAAGCACTTCAATGGTTGGCCACCGAGTCATCTTT
TACGTATGGTGGACGACGTCTCCAGGATGCTTTGATAGAGCTGGGCCCTCTGCGCTCC
TTCAAAGTGTGTTGAGGTCAAGCCTGAGAGGAGGTGGCAGGACGTGAGCATGGTGCGCATG
AAGACCATCGGGGAGCAGATCGTGGCCACATCCAGCGTGAGGTTGACTTCTCTCTG
ATGGACGTGGACAGGTCTTCAAGACGAGTTCGGGGTGGAGACCTGGGTGAGTCGGTG

CCCCAGCTACAGGCCTGGTGGTACAAGGCAGATCCCGATGAGTTTACCTACGAGAGGCGC
AAGGAGTCTGCAGCATACATTCCCTTCGGCGAAGGGGATTTTATTACCACGCAGCCATT
TTTGGGGGAACACCCACTCAGGTCCTTAACATCACCCAGGAATGCTTCAAAGGAATCCTC
AAGGACAAGAAAAATGACATAGAAGCCCAATGGCATGATGAGAGCCATCTAAACAAGTAT
TTCCTTCTCAACAAACCCACTAAAATCTTATCCCCGAATACTGCTGGGATTATCATATA
GGCCTACCTGCGGATATTAAGCTTGTCAAGATGTCTTGGCAGACAAAAGAGTATAATGTG
GTTAGAAATAACGTCTGA
[end of coding sequence]

SEQ. ID NO:7

MNVKGVILS MLVSTVIVV FWEYIHSPEG SLFWINPSRN PEVSGGSSIQ
KGWFFPRWFN NGYQEEDVD DEEKEQRKED KSKLKLSDWF NPFKRPEVVT
MTDWKAPVWV EGTYNRAVLDDYYAKQKITV GLTVFAVGRI IEHYLEEFLLT
SANKHFMVGH RVIFYVMVDD VSRMPLIELG PLRSFKVFEV KPERRWQDVS
MVRMKTIGEHI IVAHIQREVD FLFCMDVDQV FQDEFQVETL GESVAQLQAW
WYKADPDEFT YERRKESAAY IPFGEGDFY HAAIFGGTPT QVLNITQECF
KGILKDKKND IEAQWHDSEH LNKYFLLNKP TKILSPEYCW DYHIGLPADI
KLVKMSWQTK EYNVVRNNV*

SEQ. ID NO:8

1 atgaatgtca aaggaaaagt aattctgtcg atgctggttg tctcaactgt gattgtgtgtg
61 ttttgggaat atatcaacag cccagaaggc tcttcttctt ggatatatca ctcaaagaac
121 ccagaagttg atgacagcag tgctcagaag gactggtggt ttcctggctg gtttaacaat
181 gggatccaca attatcaaca agaggaagaa gacacagaca aagaaaaagg aagagaggag
241 gaacaaaaaa aggaagatga cacaacagag cttcggctat gggactggtt taatccaaaag
301 aaacgcccag aggttatgac agtgacccaa tggaggcgc cggttgtgtg ggaaggcact
361 tacaacaaag ccattcctaga aaattattat gccaaacaga aaattaccgt ggggttgacg
421 gtttttgcta ttggaagata tattgagcat tacttgagg agttcgtaac atctgcta
481 aggtacttca tggctcgcca caaagtcata tttatgtca tggtgatga tgtctccaag
541 gcgcccgtta tagagctggg tcctctgctg tccttcaaag tgtttgaggt caagccagag
601 aagaggtggc aagacatcag catgatgctg atgaagacca tcggggagca catcttgcc
661 caccatccaac acgaggttga cttcctcttc tgcattggtg tggaccagggt cttccaagac
721 cattttgggg tagagaccct gggccagtcg gtggtcagc tacaggcctg gtgttacaag
781 gcagatcctg atgactttac ctatgagagg cggaaagagt cggcagcata tattccattt
841 ggccaggggg atttttatta ccatgcagcc atttttggag gaacaccgat tcagggtctc
901 aacatcacc aggagtgtt taagggaatc ctctggaca agaaaaatga catagaagcc
961 gagtggcatg atgaaagcca cctaaacaag tatttcttc tcaacaacc ctctaaatc
1021 ttatctccag aatactgctg ggattatcat ataggcctgc cttcagatat taaaactgtc
1081 aagctatcat ggcaacaaa agagtataat ttggttagaa agaattgtctg a

SEQ. ID NO:9

MNVKGVILSMLVSTVIVVFWYINSPEGSFLWIYHKNPEVD
DSSAQKDWFPWFNNGIHNYYQEEEDTDKEKGREEEQKEDDTTELRLWDWFPNPKR
PEVMTVTQWKAPVWVWEGTYNKAILENYYAKQKITVGLTVFAIGRYIEHYLEEFVTSAN
RYFMVGHKVIIFYVMVDDVSKAPFIELGPLRSFKVFEVKEPKRWQDISMMRMKTIGEHI
LAHIQHEVDFLFCMDVDQVFDHFQVETLQSSAQLQAWWYKADPDFTYERRKESAA
YIPFGQGFYHAAIFGGTPIQVLNITQECFKGILLDKKNDIEAEWHDSEH LNKYFLL
NKPSKILSPEYCWYHIGLPSDIKTVKLSWQTKENLVRNNV

SEQ. ID NO:10

1 atggccgagg tgttgcggac gctggccgga aaacaaaaat gccacgcact tcgacctatg
61 atccttttcc taataatgct tgtcttggc ttgtttggtt acgggggtcct aagccccaga
121 agtctaatagc caggaagcct ggaacggggg ttctgcatgg ctgttaggga acctgacctat
181 ctgcagcgcg tctctgttgc aaggatgtgc taccctcagc caaagggtgct gacaccgtg
241 aaggatgtcc tcgtggtgac cccttggtg gctccattg tctgggaggg cacattcaac
301 atcgacatcc tcaacgagca gttcaggctc cagaacacca ccattgggtt aactgtgtt
361 gccatcaaga aatacgtggc tttctgaag ctgttctggt agacggcgga gaagcacttc
421 atggtgggccc accgtgtcca ctactatgtc ttcaccgacc agctggccgc ggtgccccgc
481 gtgacgctgg ggaccggctg gcagctgtca gtgctggagg tgcgcgccta caagcgtctg
541 caggacgtgt ccatgcgccc catggagatg atcagtgtg tctgcgagcg gcgcttccctc

601 agcgagggtg attacctggt gtgcgtggac gtggacatgg agttccgcga ccacgtgggc
661 gtggagatcc tgactccgct gttcggcacc ctgcaccccg gcttctacgg aagcagccgg
721 gaggccttca cctacgagcg ccggccccag tcccaggcct acatcccca aagcaggggc
781 gatttctact acctgggggg gttcttcggg gggtcgggtg aagagggtga gcggctcacc
841 agggcctgcc accaggccat gatggtcgac caggccaacg gcacgaggc cgtgtggcac
901 gacgagagcc acctgaacaa gtacctgctg cgcacacaac ccaccaagg gctctcccc
961 gagtacttgt gggaccagca gctgctgggc tggcccccg tcctgaggaa gctgagggtc
1021 actgcggtgc ccaagaacca ccaggcggtc cggaaccggt ga

SEQ. ID NO:11

MAEVLRTLAGPKCHALRPMILFLIMLVLFYGVLSPRSLMP
GSLERGFMAVREPDHLQRVSLPRMVYPQPKVLTWPKDVLVTPWLAPIVWEGTFNID
ILNEQFRLQNTTIGLTVFAIKKYVAFKLFLTAEKHFVGHVRVHYVFTDQLAAVPR
VTLGTGRQLSVLEVAYKRWQDVSMRRMEMISDFCERRFLSEVDYLVCDVDMEFRDH
VGVEILTPLFGLHPGFYGSREAFYERRPQSQAYIPKDEGDFYLLGGFFGGSVQEV
QRLTRACHQAMMVDQANGIEAVWHDESHLNKYLRLHKPTKVLSPYLDQQLLQWPAV
LRKLRFATVPKNHQAVRNP

SEQ. ID NO:12

1 atggccgagg tgttgcggac gctggccgga aaacaaaaat gccacgcact tcgacctatg
61 atccttttcc taataatgct tgtcttgggt ttgtttggtt acgggggtcct aagccccaga
121 agtctaagtc caggaagcct ggaacggggg ttctgcatgg ctgttaggga acctgacct
181 ctgcagcgcg tctcgttgcc aaggatgggt tacccccagc caaagggtgct gacaccgtgt
241 aggaaggatg tcctcgtggt gaccccttgg ctggctccca ttgtctggga gggcacgttc
301 aacatcgaca tcctcaacga gcagttcagg ctccagaaca ccaccattgg gttactgtg
361 tttgccatca agaaatacgt ggctttcctg aagctgttcc tggagacggc ggagaagcac
421 ttcattggtg gccaccgtgt ccactactat gtcttcaccg accagccggc cgcggtgccc
481 cgcgtgacgc tggggaccgg tcggcagctg tcagtgctgg aggtgggccc ctacaagcgc
541 tggcaggacg tgtccatgcg ccgcatggag atgatcagtg acttctgcga gcggcgcttc
601 ctccagcagg tggattacct ggtgtgcgtg gacgtggaca tggagtccg cgacctgtg
661 ggcgtggaga tcctgactcc gctgttcggc accctgcacc ccagcttcta cggaaagcgc
721 cgggaggcct tcacctacga gcgccggccc cagtcccagg cctacatccc caaggacgag
781 ggcgatttct actacatggg ggcgttcttc ggggggtcgg tgcaagaggt gcagcggtc
841 accagggcct gccaccaggc catgatgggt gaccaggcca acggcatcga ggccgtgtg
901 cacgacgaga gccacctgaa caagtaccta ctgcgcaca aaccaccaa ggtgtcttc
961 cccgagtact tgtgggacca gcagctgctg ggctggccc cgcctctgag gaagctgagg
1021 ttcactgcgg tgcccaagaa ccaccaggcg gtccggaacc cgtga

SEQ. ID NO:13

MAEVLRTLAGPKCHALRPMILFLIMLVLFYGVLSPRSLMP
GSLERGFMAVREPDHLQRVSLPRMVYPQPKVLTWPKDVLVTPWLAPIVWEGTFNI
DILNEQFRLQNTTIGLTVFAIKKYVAFKLFLTAEKHFVGHVRVHYVFTDQPAVPR
RVTLGTGRQLSVLEVAYKRWQDVSMRRMEMISDFCERRFLSEVDYLVCDVDMEFRDH
HVGVEILTPLFGLHPGFYGSREAFYERRPQSQAYIPKDEGDFYLLGGFFGGSVQEV
VQRLTRACHQAMMVDQANGIEAVWHDESHLNKYLRLHKPTKVLSPYLDQQLLQWPAV
VRKLRFATVPKNHQAVRNP

SEQUENCE LISTING

<110> Geron Corporation
McWhir, Jim
Gold, Joseph D.
Schiff, J. Michael

<120> 096,004 - SeqList

<130> 096,004 - SeqList

<140> [to be assigned]

<141> 2001-11-26

<150> 60/253,357

<151> 2000-11-27

<160> 20

<170> PatentIn version 3.1

<210> 1

<211> 15418

<212> DNA

<213> Homo sapiens

<400> 1

gcggccgcga	gctctaatac	gactcactat	agggcgctcg	ctcgatcaat	ggaagatgag	60
gcattgccga	agaaaagatt	aatggatttg	aacacacagc	aacagaaact	acatgaagtg	120
aaacacagga	aaaaaaagat	aaagaaacga	aaagaaaagg	gcatcagtga	gcttcagcag	180
aagttccatc	ggccttacat	atgtgtaagc	agaggccctg	taggagcaga	ggcaggggga	240
aaatacttta	agaaaataatg	tctaaaagtt	tttcaaatac	gaggaaaaac	ataaaaccac	300
agatccaaga	agctcaacaa	aacaaagcac	aagaaacagg	aagaaattaa	aagttatatac	360
acagtcaaat	tgctgaaaac	cagcaacaaa	gagaatatct	taagagtatac	agaggaaaag	420
agattaatga	caggccaaga	aacaatgaaa	acaatacaga	tttcttgtag	gaaacacaag	480
acaaaagaca	ttttttaaaa	cctaaaaggaa	aaaaaatgct	acattaaaaat	gttttttacc	540
cactgaaagt	atatttcaaa	acatatattta	ggccaggctt	ggtgggtcac	acctgtaatac	600
ccagcacttt	gggaggccaa	ggtgggtgga	tcgcttaagg	tcaggagttc	gagaccagcc	660
tggccaatat	agcgaacccc	catctgtact	aaaaacacaa	aaattagctg	ggtgtggtga	720
cacatgcctg	taatcccagg	tactcaggag	gctaaggcag	gagaattgct	tgaactggga	780
ggcagagggt	gtgagccaag	attgcaccag	tgactccag	ccttggtgac	agagtgaaac	840
tccatctcaa	aaacaaacaa	acaaaataca	tatacatata	tatatatgca	catatatata	900
catatatataa	tatatataca	catatatataa	tctatatata	tatatacata	tatacacata	960
tataaatcta	tatacatata	tatacatata	taatatattt	acatatataa	atataatcat	1020
atataaatat	acatatataa	atatacatat	aaatatacat	atataaatat	acatatataa	1080
atatacatat	ataaatatat	atatacatat	ataaatatat	ataaatatat	atacatatat	1140
aaatatataa	atatacaagt	atatacaaat	atatacatat	ataaatgtat	atacgtatat	1200
acatatatat	ataaatatat	aaaaaaactt	ttggctgggc	acctttccaa	atctcatggc	1260
acatatatagt	ctcatggtaa	cctcaaatata	aaaaacatat	aacagatata	ccaaaaataa	1320
aaaccaataa	attaaatcat	gccaccagaa	gaaattacct	tcactaaaag	gaacacagga	1380
aggaaaagaaa	gaagggaagag	aagaccatga	aacaaccaga	aaacaaacaa	caaaacagca	1440
ggagtaattc	ctgacttatac	aataataatg	ctgggtgtaa	atggactaaa	ctctccaatc	1500
aaaagacata	gagtggctga	atggacgaaa	aaaacaagac	tcaataatct	gttgcttaca	1560
agaatatact	tcacctataa	agggacacat	agactgaaaa	taaaagggaag	gaaaaatatt	1620
ctatgcaaat	ggaaacccaaa	aaaagaacag	aactagctac	acttatatca	gacaaaatag	1680
atttcaagac	aaaaagtata	aaaagagaca	aagtaattat	ataataataa	agcaaaaaga	1740
tataacaatt	gtgaatttat	atgcgcccac	cactgggaca	cccagatata	tacagcaaat	1800
attattagaa	ctaaggagag	agagagatcc	ccatacaata	atagctggag	acttcacccc	1860
gcttttagca	tgggacagat	catccagaca	gaaaatcaac	caaaaaattg	gacttaatact	1920
ataatataga	acaaatgtac	ctaattgatg	tttacaagac	atttcatcca	gtagttgcag	1980
aatatgcatt	ttttctcag	catatggatc	attctcaagg	atagaccata	tattaggcca	2040
cagaacaagc	cattaaaaat	tcaaaaaaat	tgagccaggc	atgatggctt	atgcttgtaa	2100
ttacagcact	ttggggaggg	tgaggtggga	ggatgtcttg	agtacaggag	tttgagacca	2160
gcctgggcaa	aatagtgaga	ccctgtctct	acaaactttt	tttttaatt	agccaggcat	2220
agtgggtgtg	gcctgtagtc	ccagctactt	aggaggctga	agtgggagga	tcacttgagc	2280

ccaagagttc	aaggctacgg	tgagccatga	ttgcaacacc	acacaccagc	cttggtgaca	2340
gaatgagacc	ctgtctcaaa	aaaaaaaaaa	aaaattgaaa	taataataag	catcttctct	2400
ggccacagtg	gaacaaaacc	agaaatcaac	aacaagagga	attttgaaaa	ctatacaaac	2460
acatgaaaa	taacaatat	acttctgaat	aaccagttag	tcaatgaaga	aattaaaaag	2520
gaaattgaaa	aattttattt	agcaaatgat	aacggaaaaca	taacctctca	aaacccacgg	2580
tatacagcaa	aagcagtgtc	aagaaggaag	tttatagcta	taagcagcta	catcaaaaaa	2640
gtagaaaagc	caggcgagtg	ggctcatgcc	tgtaatccca	gcactttggg	aggccaagcg	2700
gggcagatcg	cctgagggtca	ggagttcgag	accagcctga	ccaacacaga	gaaaccttgt	2760
cgctactaaa	aatacaaaat	tagctgggca	tggtggcaca	tgcttgaat	cccagctact	2820
cgggaggctg	aggcaggata	accgcttgaa	cccaggaggt	ggagggttgcg	gtgagccggg	2880
attgcgccat	tggactccag	cctgggtaac	aagagtgaag	ccctgtctca	agaaaaaaa	2940
aaaagttaga	aaacttaaaa	atacaacct	atgatgcacc	ttaaagaact	agaaaaagca	3000
gagcaaaata	aacctaaaat	tggtaaaaga	aaagaaataa	taaagatcag	agcagaataa	3060
aatgaaactg	aaagataaca	atacaaaaga	tcaacaaaat	taaaagttag	ttttttgaaa	3120
agataaaaca	aattgacaaa	cctttgccca	gactaagaaa	aaaggaaaga	agacctaaat	3180
aaataaagtc	agagatgaaa	aaagagacat	tacaactgat	accacagaaa	ttcaaaggat	3240
cactagaggg	tactatgagc	aactgtacac	taataaattg	aaaaacctag	aaaaaataga	3300
taaatctcta	gatgcataca	acctaccaag	attgaacat	gaagaaatcc	aaagcccaaa	3360
cagaccaata	acaataatgg	gattaaagcc	ataataaaaa	gtctcctagc	aaagagaagc	3420
ccaggaccca	atggcttccc	tgctggattt	taccaatcat	ttaaagaaga	atgaattcca	3480
atcctactca	aactattctg	aaaaatagag	gaaagaatac	ttccaaactc	attctacatg	3540
gccagtatta	ccctgatttc	aaaaccagac	aaaaacacat	caaaaaacaa	caaacaaaaa	3600
aacagaaaga	aagaaaacta	caggccaata	tccttgatga	atactgatac	aaaaatcctc	3660
aacaaaaac	tagcaaacca	aattaaacaa	cacttctgaa	agatcattca	ttgtgatcaa	3720
gtgggattta	ttccagggat	ggaaggatgg	ttcaacatat	gcaaatcaat	caatgtgata	3780
catcatccca	acaaaatgaa	gtacaaaaac	tatatgatta	ttcacttta	tgcagaaaaa	3840
gcatttgata	aaattctgca	cccttcatga	taaaaacctt	caaaaaacca	ggtatacaag	3900
aaacatacag	gccaggcaca	gtggctcaca	cctgcgatcc	cagcactctg	ggaggccaag	3960
gtgggatgat	tgcttggggc	caggagttag	agactagcct	gggcaacaaa	atgagacctg	4020
gtctacaaaa	aactttttta	aaaaattagc	caggcatgat	ggcatatgcc	tgtagtcccc	4080
gctagtctgg	aggctgaggt	cttgaagaat	cttaagccta	ggaggctgag	gctgagtgga	4140
gccatgaaca	tgctactgta	ctccagccta	gacaacagaa	caagacccca	ctgaataaga	4200
agaaggagaa	ggagaaggga	gaaaggaggg	agaaggaggg	aggaggagaa	ggaggagggtg	4260
gaggagaaat	gggaagggga	ggggaaggga	aagagggaaga	agaagaaaca	tatttcaaca	4320
taataaaaagc	cctatatgac	agaccgaggt	agtattatga	ggaaaaactg	aaagcctttc	4380
ctctaagatc	tggaaaatga	caaggggcca	ctttcaccac	tgtgattcaa	catagtacta	4440
gaagtccatg	ctagagcaat	cagataagag	aaagaaataa	aaggcatcca	aactggaaag	4500
gaagaagtca	aattatcctg	tttgagatg	atatgatctt	atatctggaa	aagacttaag	4560
acaccactaa	aaaactatta	gagctgaaat	ttggtacagc	aggatacaaa	atcaatgtac	4620
aaaaatcagt	agtatttcta	tattccaaca	gcaaaacaat	tgaaaaagaa	accaaaaaag	4680
cagctacaaa	taaaattaaa	cagctaggaa	ttaaaccaag	aagtgaaga	tctctacaat	4740
gaaaactata	aaatattgat	aaaagaaatt	gaagagggca	caaaaaaaga	aaagatattc	4800
catgttcata	gattggaaga	ataaatactg	ttaaaatgtc	catactaccc	aaagcaattt	4860
acaaattcaa	tgcaatccct	attaaaatac	taatgacgtt	cttcacagaa	atagaagaaa	4920
caatttcaag	attgttacag	aaccacaaaa	gaccagaat	agccaaaagc	atcctgacca	4980
aaaagaacaa	aactggaagc	atcacattac	ctgacttcaa	attatactac	aaagctatag	5040
taacccaac	tacatggtac	tggcataaaa	acagatgaga	catggaccag	aggaacagaa	5100
tagagaatcc	agaaacaaat	ccatgcatct	acagtgaact	catttttgac	aaagggtgca	5160
agaacatact	ttggggaaaa	gataatctct	tcaataaatg	gtgctggagg	aactggatat	5220
ccatatgcaa	aataacaata	ctagaactct	gtctctcacc	atatacaaaa	gcaaatcaaa	5280
atggatgaaa	ggcttaaatc	taaaacctca	aactttgcaa	ctactaaaag	aaaacaccgg	5340
agaaactctc	caggacattg	gagtgggcaa	agacttcttg	agtaattccc	tgcaaggcaca	5400
ggcaacccaa	gcaaaaacag	acaaatggga	tcatatcaag	ttaaaaagct	tctgccagc	5460
aaaggaaaca	atcaacaaag	agaagagaca	accacagaa	tgaggagaata	tatttgcaaa	5520
ctattcatct	aacaaggaat	taataaccag	tatatataag	gagctcaaac	tactctataa	5580
gaaaaacacc	taataagctg	attttcaaaa	ataagcaaaa	gatctgggta	gacatttctc	5640
aaaataagtc	atacaaatgg	caaacaggca	tctgaaaatg	tgctcaacac	cactgatcat	5700
cagagaatag	caaatcaaaa	ctactatgag	agatcatctc	accccagtta	aaatggcttt	5760
tattcaaaag	acaggcaata	acaaatggca	gtgaggatgt	ggataaaaag	aaacccttgg	5820
acactgtttg	tggggaatgga	aattgctacc	actatggaga	acagtttgaa	agttcctcaa	5880
aaaactaaaa	ataaagctac	catacagcaa	tcccattgct	aggtatatac	tccaaaaaag	5940
ggaatcagtg	tatcaacaag	ctatctccac	tcccacattt	actgcagcac	tgttcatagc	6000
agccaaggtt	tggaagcaac	ctcagtgtcc	atcaacagac	gaatggaaaa	agaaaaatgtg	6060
gtgcacatac	acaatggagtg	actacgcagc	cataaaaaag	aatgagatcc	tgctcagttgc	6120
aacagcatgg	ggggcactgg	tcagtatggt	aagtgaataa	agccaggcac	agaaagacaa	6180
acttttcatg	ttctccctta	cttgtggggg	caaaaattaa	aacaattgac	atagaataatg	6240
aggagaatgg	tggttctaga	gggggtgggg	acagggtgac	tagagtcaac	aataatttat	6300
tgtatgtttt	aaaataacta	aaagagtata	attgggttgt	ttgtaacaca	aagaaaggat	6360

aaatgcttga	aggtgacaga	taccccat	accctgatgt	gattattaca	cattgtatgc	6420
ctgtatcaaa	atatctcatg	tatgctatag	atataaaccc	tactatatta	aaaattaaaa	6480
ttttaattggc	caggcacggt	ggctcatgtc	cataatccca	gcactttggg	aggccgaggc	6540
ggtggatcac	ctgagggtcag	gagtttgaaa	ccagtctggc	caccatgatg	aaacctgtc	6600
tctactaaag	atacaaaaat	tagccaggcg	tggtggcaca	tacctgtagt	cccaactact	6660
caggaggctg	agacaggaga	attgcttgaa	cctgggaggc	ggagggttga	gtgagccgag	6720
atcatgccc	tgactgcag	cctgggtgac	agagcaagac	tccatctcaa	aacaaaaaca	6780
aaaaaaagaa	gattaaaaat	gtaattttta	tgtaaccgtat	aaatatatac	tctactatat	6840
tagaagttaa	aaattaaaac	aattataaaa	ggtaattaac	cacttaatct	aaaataagaa	6900
caatgtatgt	gggttttcta	gcttctgaag	aagtaaaagt	tatggccacg	atggcagaaa	6960
tgtgaggagg	gaacagtggg	agttactggt	gttagacgct	catactctct	gtaagtgtact	7020
taatttttaac	caaagacagg	ctgggagaag	ttaaagaggc	attctataag	ccctaaaaca	7080
actgctaata	atgggtgaaag	gtaatctcta	ttaattacca	ataattacag	atatctctaa	7140
aatcgagctg	cagaattggc	acgtctgatc	acaccgtcct	ctcattcacg	gtgctttttt	7200
tcttgtgtgc	ttggagattt	tcgattgtgt	gttcgtgttt	ggtaaacctt	aatctgtatg	7260
aatcctgaaa	cgaaaaatgg	tggtgatttc	ctccagaaga	attagagtac	ctggcaggaa	7320
gcagggtggct	ctgtggacct	gagccacttc	aatcttcaag	ggctctctggc	caagaccag	7380
gtgcaaggca	gaggcctgat	gacccgagga	caggaaagct	cggatgggaa	ggggcgatga	7440
gaagcctgcc	tcgttgggtga	gcagcgcag	aagtgcctct	atttacgctt	tgcaaaagatt	7500
gctctggata	ccatctggaa	aaggcgccca	gcgggaatgc	aaggagtacg	aagcctcctg	7560
ctcaaaccca	ggccagcagc	tatggcgccc	acccgggcgt	gtgccagagg	gagaggagtc	7620
aaggcacctc	gaagtatggc	tttaaatcttt	ttttcacctg	aagcagtgac	caagggtgtat	7680
tctgagggaa	gcttgagtta	gggtgccttct	ttaaaacaga	aagtcatgga	agcacccttc	7740
tcaagggaaa	accagacgcc	cgctctgcgg	tcattttacct	ctttcctctc	tccctctctt	7800
gccctcgagg	tttctgatcg	ggacagagtg	acccccgtgg	agcttctccg	agccccgtgt	7860
gaggacccctc	ttgcaaaagg	ctccacagac	ccccgcctg	gagagaggag	tctgagcctg	7920
gcttaataac	aaactgggat	gtggctgggg	gcggacagcg	acggcgggat	tcaaaagactt	7980
aattccatga	gtaaaattcaa	cctttccaca	tccgaatgga	tttggatttt	atcttaatat	8040
tttcttaaat	ttcatcaaat	aacattcagg	agtcagaaaa	tccaaaggcg	taaaacagga	8100
actgagctat	gtttgccaag	gtccaaggac	ttaataacca	tgttcagagg	gatttttctgc	8160
cctaagtact	ttttatttgt	tttcataaag	tggtcttaggg	tgcaagggaa	agtacacgag	8220
gagaggactg	ggcggcaggg	ctatgagcac	ggcaaggcca	ccggggagag	agtccccggc	8280
ctgggaggct	gacagcagga	ccactgaccg	tcctccctgg	gagctgccac	attgggcaac	8340
gcgaaggcgg	ccacgctgcg	tgtgactcag	gaccccatat	cggtctcctg	ggccccacca	8400
cactaaccca	ggaagtcacg	gagctctgaa	cccgtggaaa	cgaaactgac	ccttgccctgc	8460
ctgctttccct	gggtgggtca	aggttaatga	agtgtgtgtg	aggaaaaggc	catgtaaatt	8520
acacgactct	gctgatgggg	accgttccct	ccatcattat	tcactctcac	ccccaggac	8580
tgaatgattc	cagcaacttc	ttcgggtgtg	acaagccatg	acaacactca	gtacaacac	8640
cactctttta	ctaggccac	agagcacggc	ccacaccctt	gatatattaa	gagtccagga	8700
gagatgaggc	tgctttcagc	caccaggctg	gggtgacaac	agcggctgaa	cagtctgttc	8760
ctctagacta	gtagaccctg	gcaggcactc	ccccagattc	tagggcctgg	ttgtgtcttc	8820
ccgagggcgc	catctgccct	ggagactcag	cctgggtgtg	cacactgagg	ccagccctgt	8880
ctccacaccc	tccgcctcca	ggcctcagct	tctccagcag	cttcctaaac	cctgggtggg	8940
ccgtgttcca	gcgctactgt	ctcacctgtc	ccactgtgtc	ttgtctcagc	gacgtagctc	9000
gcacgtttcc	tctcacatg	gggtgtctgt	ctccttcccc	aacactcaca	tgctgtgaag	9060
ggaggagatt	ctgcgccttc	cagactggct	cctctgagcc	tgaacctggc	tcgtggcccc	9120
cgatgaggtt	tcctggcgct	cggctgcacg	ctgacctcca	tttccaggcg	ctccccgtct	9180
cctgtcatct	gccggggcct	gccggtgtgt	tcttctgttt	ctgtgtctct	ttccacgtcc	9240
agctgcgtgt	gtctctgtcc	gctagggtct	cgggtttttt	ataggcatag	gacgggggag	9300
tggtggggcca	ggcgctctct	gggaaatgca	acatttgggt	gtgaaaagtag	gagtgccctgt	9360
cctcacctag	gtccacgggc	acaggcctgg	ggatggagcc	cccgcagggg	acccgccctt	9420
ctctgcccag	cacttttctg	ccccctcccc	tctggaacac	agagtggcag	tttccacaag	9480
cactaagcat	cctcttccca	aaagaccag	cattggcacc	cctggacatt	tgccccacag	9540
ccctgggaat	tcacgtgact	acgcacatca	tgtaacacact	ccgttccacg	accgaccccc	9600
gctgttttat	tttaatagct	acaaagcagg	gaaatccctg	ctaaaaatgtc	ctttaacaaa	9660
ctggttaaac	aaacgggtcc	atccgcacgg	tggacagtcc	ctcacagtga	agaggaaacat	9720
gccgtttata	aagcctgcag	gcattctcaag	ggaattacgc	tgagtcaaaa	ctgccacctc	9780
catgggatac	gtacgcaaca	tgctcaaaaa	gaaagaattt	caccccatgg	caggggagtg	9840
gttggggggg	taaggacggt	gggggcagca	gctgggggct	actgcacgca	ccttttacta	9900
aagccagttt	cctggttctg	atggtattgg	ctcagttatg	ggagactaac	cataggggag	9960
tggggatggg	ggaaccggga	ggctgtgcca	tcttgccat	gcccgagtgt	cctgggcagg	10020
ataatgctct	agagatgccc	acgtcctgat	tccccaaac	ctgtggacag	aacccgcccg	10080
gccccagggc	ctttgcaggt	gtgatctccg	tgaggacctt	gaggtctggg	atccttcggg	10140
actactgtga	ggcccgaaaa	gtaatccagg	ggttctggga	agaggcgggc	aggagggtca	10200
gaggggggca	gcctcaggac	gatggaggca	gtcagtctga	ggctgaaaag	ggaggggagg	10260
cctcgagccc	aggcctgcaa	gcgcctccag	aagctggaaa	aagcggggaa	gggacctctc	10320
acggagcctg	cagcaggaag	gcacggctgg	cccttagccc	accaggggcc	atcgtggacc	10380
tccggcctcc	gtgccatagg	agggcactcg	cgctgccctt	ctagcatgaa	gtgtgtgggg	10440

atttgcagaa	gcaacaggaa	acccatgcac	tgtgaatcta	ggattatttc	aaaacaaagg	10500
tttacagaaa	catccaagga	cagggctgaa	gtgcctccgg	gcaagggcag	ggcaggcacg	10560
agtgatttta	tttagctatt	ttattttatt	tacttacttt	ctgagacaga	gttatgctct	10620
tgttgcccag	gctggagtgc	agcggcatga	tcttggtcca	ctgcaacctc	cgtctcctgg	10680
gttcaagcaa	ttctcgtgcc	tcagcctccc	aagtagctgg	gatttccaggc	gtgcaccacc	10740
acacccggct	aattttgtat	ttttagtaga	gatgggcttt	caccatgttg	gtcaggctga	10800
tctcaaaatc	ctgacctcag	gtgatccgcc	cacctcagcc	tcccaaagtg	ctgggattac	10860
aggcatgagc	cactgcacct	ggcctattta	accattttaa	aacttccctg	ggctcaagtc	10920
acacccactg	gtaaggagtt	catggagtgc	aatttcccct	ttactcagga	gttaccctcc	10980
tttgatattt	tctgtaattc	ttcgtagact	ggggatacac	cgtctcttga	catattcaca	11040
gtttctgtga	ccacctgtta	tcccatggga	cccactgcag	gggcagctgg	gaggctgcag	11100
gcttcaggtc	ccagtggggg	tgccatctgc	cagtagaaaac	ctgatgtaga	atcaggggcg	11160
gagtgtggac	actgtcctga	atctcaatgt	ctcagtgtgt	gctgaaacat	gtagaaatta	11220
aagtccatcc	ctcctactct	actgggattg	agccccctcc	ctatcccccc	ccaggggcag	11280
aggagttcct	ctcactcctg	tggaggaagg	aatgatactt	tgttattttt	cactgtctgg	11340
actgaatcca	ctggttcatt	tgttggtttg	tttgttttgt	tttgagaggc	ggtttcactc	11400
ttgttgctca	ggctggaggg	agtgaatgg	cgcatctctg	gcttactgca	gcctctgcct	11460
cccagggtta	agtgattctc	ctgcttccgc	ctcccatttg	gctgggatta	caggcacccg	11520
ccacatgccc	cagctaattt	tttgattttt	tagtagagac	gggggtgggg	gtggggttca	11580
ccatgttggc	caggctgggc	tcgaacttct	gacctcagat	gatccacctg	cctctgcctc	11640
ctaaagtgtc	gggattacag	gtgtgagcca	ccatgcccag	ctcagaattt	actctgttta	11700
gaaacatctc	ggctctgagg	aggaagctca	ccccactcaa	gtgtgtgtgt	gttttaagcc	11760
aatgatagaa	tttttttatt	gttgttagaa	cactcttgat	gttttacct	gtgatgacta	11820
agacatcatc	agcttttcaa	agacacacta	actgcaccca	taatactggg	gtgtctctctg	11880
ggtatcagcg	atcttcattg	aatgccggga	ggcgtttcct	cgccatgcac	atggtgttaa	11940
ttactccagc	ataatcttct	gcttccattt	cttctcttcc	ctcttttaaa	attgtgtttt	12000
ctatgttggc	ttctctgcag	agaaccagtg	taagctacaa	cttaactttt	gttggaacaa	12060
attttccaaa	ccgccccctt	gccctagtgg	cagagacaat	tcacaaacac	agccccctta	12120
aaaggcttag	ggatcactaa	ggggatttct	agaagagcga	cccgaatcc	taagtattta	12180
caagacgagg	ctaacttcca	gcgagcgtga	cagcccaggg	agggtgcgag	gcctgttcaa	12240
atgctagctc	cataaataaa	gcaatttctc	ccggcagttt	ctgaaaagtag	gaaaggttac	12300
atttaaggtt	gcgtttgtta	gcatttcagt	gtttgccgac	ctcagctaca	gcatccctgc	12360
aaggcctcgg	gagaccagaa	agtttctcgc	cccttagatc	caaacttgag	caaccgggag	12420
tcttgattcc	tgggaagtcc	tcagctgtcc	tgcggttgtg	ccggggcccc	aggctctggag	12480
gggaccagtg	gccgtgtggc	ttctactgtg	gggctggaag	tcgggctctc	tagctctgca	12540
gtccgaggct	tggagccagg	tgcgtgtacc	ccgaggctgc	cctccacctc	gtgcgggcgg	12600
gatgtgacca	gatgttgccc	tcattctgcca	gacagagtgc	cggggcccag	ggtcaaggcc	12660
gttgtgctg	gtgtgaggcg	cccgtgtcgc	ggccagcagg	agcgcctggc	tccatttccc	12720
acccctttctc	gacgggaccg	ccccgggtgg	tgattaacag	atttgggggtg	gtttgtctat	12780
gggtggggac	cctgcgcgcc	tcgaacacct	caaagagaaa	tgacgggcct	gtgtcaagga	12840
gccccagtgc	cggggaaagt	ttgcaggagg	gacttccggg	aggctccggc	tgcccgtcca	12900
gggagcaatg	cgctctcggg	ttcgtcccca	gccgcgtcta	cgccctccg	tcctccccct	12960
cacgtccggc	attcgtgggt	cccggagccc	gacgccccgc	gtccggacct	ggaggcagcc	13020
ctgggtctcc	ggatcaggcc	agcggccaaa	gggtcgccgc	acgcacctgt	tcccaggggc	13080
ttcacatcat	ggccccctcc	tcgggttacc	ccacagccta	ggccgattcg	acctctctcc	13140
gctggggccc	tcgctggcgt	ccctgcaccc	tgggagcgcg	agcggcgcg	gggaggggaa	13200
gcgcggccca	gacccccggg	tccgcccggg	gcagctgcgc	tgtcggggcc	aggccgggct	13260
cccagtggtg	tcgcgggcac	agacgcccag	gaccgcgttt	cccacgtggc	ggagggactg	13320
gggaccggcg	caccgcctcc	gccccctcac	cttcagctc	cgctctctcc	gcgcggaccc	13380
cgccccgtcc	cgacccccct	cgggttcccc	gccccagccc	ctccggggcc	tcccagcccc	13440
tccccctctc	ttccgcggcc	ccgccccctc	ctcgggcgcg	gagtttcagg	cagcgtgcgc	13500
tcctgctgcg	cacgtgggaa	gccccggccc	cgcccacccc	cgcatgccc	cgctctcccc	13560
gctgccgagc	cgctgcctcc	ctgctgcgca	gccactaccg	cagggtgctg	ccgctggcca	13620
cgcttcgtgc	gcgcctgggg	ccccagggct	ggcggtgggt	gcagcgcggg	gaccggcggg	13680
ctttccgcgc	gctggtggcc	cagtgctctg	tgtgcgtgct	ctgggacgca	cggccggccc	13740
ccgcccggcc	ctccttccgc	cagggtggcc	tccccggggt	cggcgctcgg	ctgggggtga	13800
gggcggccgg	ggggaaccag	cgacatgcgg	agagcagcgc	aggcgactca	gggcgcttcc	13860
cccgcagggt	tcctgcctga	aggagctggt	ggcccagagt	ctgcagaggc	tgtgcgagcg	13920
cggcgcgaag	aacgtgctgg	ccttcggctt	cgcgctgctg	gacggggccc	gcggggggcc	13980
ccccgaggcc	ttcaccacca	gcgtgcgcag	ctacctgccc	aacacggtga	ccgacgcact	14040
gcggggggagc	ggggcggtgg	ggctgctgct	gcgcgcgcgt	ggcgacgacg	tgctggttca	14100
cctgctggga	cgctgcgcgc	tctttgtgct	ggtggctccc	agctgcgcct	accagggtgtg	14160
cgggcccggc	ctgtaccagc	tcggcgctgc	cactcaggcc	cggccccccg	cacacgctag	14220
tggaccgccga	agggctctgg	gatgcgaacg	ggcctggaac	catagcgtca	gggaggccgg	14280
ggtccccctg	ggcctgccag	ccccgggtgc	gaggaggcgc	gggggcagtg	ccagccgaag	14340
tctgccgttg	cccaagaggc	ccaggcgctg	cgctgccccct	gagccggagc	ggacgcccgt	14400
tgggcagggg	tcctggggcc	acccgggcag	gacgcgtgga	ccgagtgacc	gtggtttctg	14460
tgtggtgtca	cctgccagac	ccgccgaaga	agccacctct	ttggagggtg	cgctctctgg	14520

SEQUENCE LISTING

cacgcgccac	tcccacccat	ccgtggggccg	ccagcaccac	gcggggccccc	catccacatc	14580
gcggccacca	cgccccctggg	acacgccttg	tcccccggtg	tacgccgaga	ccaagcactt	14640
cctctactcc	tcaggcgaca	aggagcagct	gcggccctcc	ttcctactca	gctctctgag	14700
gccccagcctg	actggcgctc	ggaggctcgt	ggagaccatc	tttctgggtt	ccaggccctg	14760
gatgccaggg	actccccgca	ggttgccccg	cctgccccag	cgctactggc	aaatgcggcc	14820
cctgtttctg	gagctgcttg	ggaaccacgc	gcagtgcctc	tacgggggtg	tcctcaagac	14880
gcactgcccc	ctgcgagctg	cggtcacccc	agcagccggt	gtctgtgccc	gggagaagcc	14940
ccagggtctt	gtggcgggcc	ccgaggagga	ggacacagac	ccccgtcgcc	tggtgcagct	15000
gctccgccag	cacagcagcc	cctggcaggt	gtacggcttc	gtgcgggcct	gcctgcgccg	15060
gctgtgtccc	ccaggcctct	ggggctccag	gcacaacgaa	cgccgcttcc	tcaggaaacac	15120
caagaagttc	atctccctgg	ggaagcatgc	caagctctcg	ctgcaggagc	tgacgtggaa	15180
gatgagcgtg	cgggactgctg	cttggtctgcg	caggagccca	ggtgaggagg	tggtggccgt	15240
cgaggggccca	ggccccagag	ctgaatgcag	taggggtctca	gaaaaggggg	caggcagagc	15300
cctgtgtctc	ctgtctccat	cgtcacgtgg	gcacacgtgg	cttttcgctc	aggacgtcga	15360
gtggacacgg	tgatcgagtc	gactcccttt	agtgagggtt	aattgagctc	gcggccgc	15418

<210> 2
<211> 25
<212> DNA
<213> Artificial

<400> 2

cttgctgcag	aagtgggtgg	aggaa	25
------------	------------	-------	----

<210> 3
<211> 21
<212> DNA
<213> Artificial

<400> 3

ctgcagtgtg	ggtttcgggc	a	21
------------	------------	---	----

<210> 4
<211> 20
<212> DNA
<213> Artificial

<400> 4

cggaagagtg	tctggagcaa	20
------------	------------	----

<210> 5
<211> 19
<212> DNA
<213> Artificial

<400> 5

ggatgaagcg	gagtctgga	19
------------	-----------	----

<210> 6
<211> 1303
<212> DNA
<213> Ovis aries

<220>
<221> CDS
<222> (194)..(1303)
<223>
<400> 6

```

agccgaggac gccgccggg agccgaggct ccggccagcc cccagcgcgc ccagcttctg 60
cagatcagga gtcagaacgc tgcaccttcg cttcctccca gccctgcctc cttctgcaaa 120
acggagctca atagaacttg gtacttttgc cttttactct gggaggagag aagcagacga 180
tgaggagaaa ata atg aat gtc aaa gga aaa gtg att ctg tca atg ctg 229
      Met Asn Val Lys Gly Lys Val Ile Leu Ser Met Leu
      1          5          10
gtt gtc tca act gtc att gtt gtg ttt tgg gaa tat atc cac agc cca 277
Val Val Ser Thr Val Ile Val Val Phe Trp Glu Tyr Ile His Ser Pro
      15          20          25
gaa ggc tct ttg ttc tgg ata aac cca tca aga aac cca gaa gtc agt 325
Glu Gly Ser Leu Phe Trp Ile Asn Pro Ser Arg Asn Pro Glu Val Ser
      30          35          40
ggc ggc agc agc att cag aag ggc tgg tgg ttt ccg aga tgg ttt aac 373
Gly Gly Ser Ser Ile Gln Lys Gly Trp Trp Phe Pro Arg Trp Phe Asn
      45          50          55          60
aat ggt tac caa gaa gaa gat gaa gac gta gac gaa gaa aag gaa caa 421
Asn Gly Tyr Gln Glu Asp Glu Asp Val Asp Glu Glu Lys Glu Gln
      65          70          75
aga aag gaa gac aaa agc aag ctt aag cta tcg gac tgg ttc aac cca 469
Arg Lys Glu Asp Lys Ser Lys Leu Lys Leu Ser Asp Trp Phe Asn Pro
      80          85          90
ttt aaa cgc cct gag gtt gtg act atg aca gat tgg aag gca ccc gtg 517
Phe Lys Arg Pro Glu Val Val Thr Met Thr Asp Trp Lys Ala Pro Val
      95          100          105
gtg tgg gaa ggc act tac aac aga gcc gtc tta gac gat tac tac gcc 565
Val Trp Glu Gly Thr Tyr Asn Arg Ala Val Leu Asp Asp Tyr Tyr Ala
      110          115          120
aag cag aaa att acc gtc ggc ctg acg gtt ttc gcc gtc gga aga tac 613
Lys Gln Lys Ile Thr Val Gly Leu Thr Val Phe Ala Val Gly Arg Tyr
      125          130          135          140
att gag cat tac ttg gag gac ttc tta acg tct gct aat aag cac ttc 661
Ile Glu His Tyr Leu Glu Glu Phe Leu Thr Ser Ala Asn Lys His Phe
      145          150          155
atg gtt ggc cac cga gtc atc ttt tac gtc atg gtg gac gac gtc tcc 709
Met Val Gly His Arg Val Ile Phe Tyr Val Met Val Asp Asp Val Ser
      160          165          170
agg atg cct ttg ata gag ctg ggc cct ctg cgc tcc ttc aaa gtg ttt 757
Arg Met Pro Leu Ile Glu Leu Gly Pro Leu Arg Ser Phe Lys Val Phe
      175          180          185
gag gtc aag cct gag agg agg tgg cag gac gtc agc atg gtg cgc atg 805
Glu Val Lys Pro Glu Arg Arg Trp Gln Asp Val Ser Met Val Arg Met
      190          195          200
aag acc atc ggg gag cac atc gtg gcc cac atc cag cgt gag gtt gac 853
Lys Thr Ile Gly Glu His Ile Val Ala His Ile Gln Arg Glu Val Asp
      205          210          215          220
ttc ctc ttc tgc atg gac gtg gac cag gtc ttc caa gac gag ttc ggg 901
Phe Leu Phe Cys Met Asp Val Asp Gln Val Phe Gln Asp Glu Phe Gly
      225          230          235
gtg gag acc ctg ggt gag tcg gtg gcc cag cta cag gcc tgg tgg tac 949
Val Glu Thr Leu Gly Glu Ser Val Ala Gln Leu Gln Ala Trp Trp Tyr
      240          245          250
aag gca gat ccc gat gag ttt acc tac gag agg cgc aag gag tct gca 997
Lys Ala Asp Pro Asp Glu Phe Thr Tyr Glu Arg Arg Lys Glu Ser Ala
      255          260          265
gca tac att ccc ttc ggc gaa ggg gat ttt tat tac cac gca gcc att 1045
Ala Tyr Ile Pro Phe Gly Glu Gly Asp Phe Tyr Tyr His Ala Ala Ile
      270          275          280
ttt ggg gga aca ccc act cag gtc ctt aac atc acc cag gaa tgc ttc 1093
Phe Gly Gly Thr Pro Gln Val Leu Asn Ile Thr Gln Glu Cys Phe
      285          290          295          300
aaa gga atc ctc aag gac aag aaa aat gac ata gaa gcc caa tgg cat 1141
Lys Gly Ile Leu Lys Asp Lys Lys Asn Asp Ile Glu Ala Gln Trp His
      305          310          315
gat gag agc cat cta aac aag tat ttc ctt ctc aac aaa ccc act aaa 1189
Asp Glu Ser His Leu Asn Lys Tyr Phe Leu Leu Asn Lys Pro Thr Lys
      320          325          330
atc tta tcc ccg gaa tac tgc tgg gat tat cat ata ggc cta cct gcg 1237
Ile Leu Ser Pro Glu Tyr Cys Trp Asp Tyr His Ile Gly Leu Pro Ala

```

335 340 345
gat att aag ctt gtc aag atg tct tgg cag aca aaa gag tat aat gtg 1285
Asp Ile Lys Leu Val Lys Met Ser Trp Gln Thr Lys Glu Tyr Asn Val
350 355 360
gtt aga aat aac gtc tga 1303
Val Arg Asn Asn Val
365

<210> 7
<211> 369
<212> PRT
<213> Ovis aries

<400> 7

Met Asn Val Lys Gly Lys Val Ile Leu Ser Met Leu Val Val Ser Thr
1 5 10 15
Val Ile Val Val Phe Trp Glu Tyr Ile His Ser Pro Glu Gly Ser Leu
20 25 30
Phe Trp Ile Asn Pro Ser Arg Asn Pro Glu Val Ser Gly Gly Ser Ser
35 40 45
Ile Gln Lys Gly Trp Trp Phe Pro Arg Trp Phe Asn Asn Gly Tyr Gln
50 55 60
Glu Glu Asp Glu Asp Val Asp Glu Glu Lys Glu Gln Arg Lys Glu Asp
65 70 75 80
Lys Ser Lys Leu Lys Leu Ser Asp Trp Phe Asn Pro Phe Lys Arg Pro
85 90 95
Glu Val Val Thr Met Thr Asp Trp Lys Ala Pro Val Val Trp Glu Gly
100 105 110
Thr Tyr Asn Arg Ala Val Leu Asp Tyr Tyr Ala Lys Gln Lys Ile
115 120 125
Thr Val Gly Leu Thr Val Phe Ala Val Gly Arg Tyr Ile Glu His Tyr
130 135 140
Leu Glu Glu Phe Leu Thr Ser Ala Asn Lys His Phe Met Val Gly His
145 150 155 160
Arg Val Ile Phe Tyr Val Met Val Asp Asp Val Ser Arg Met Pro Leu
165 170 175
Ile Glu Leu Gly Pro Leu Arg Ser Phe Lys Val Phe Glu Val Lys Pro
180 185 190
Glu Arg Arg Trp Gln Asp Val Ser Met Val Arg Met Lys Thr Ile Gly
195 200 205
Glu His Ile Val Ala His Ile Gln Arg Glu Val Asp Phe Leu Phe Cys
210 215 220
Met Asp Val Asp Gln Val Phe Gln Asp Glu Phe Gly Val Glu Thr Leu
225 230 235 240
Gly Glu Ser Val Ala Gln Leu Gln Ala Trp Trp Tyr Lys Ala Asp Pro
245 250 255
Asp Glu Phe Thr Tyr Glu Arg Arg Lys Glu Ser Ala Ala Tyr Ile Pro
260 265 270

Phe Gly Glu Gly Asp Phe Tyr Tyr His Ala Ala Ile Phe Gly Gly Thr
275 280 285
Pro Thr Gln Val Leu Asn Ile Thr Gln Glu Cys Phe Lys Gly Ile Leu
290 295 300
Lys Asp Lys Lys Asn Asp Ile Glu Ala Gln Trp His Asp Glu Ser His
305 310 315 320
Leu Asn Lys Tyr Phe Leu Leu Asn Lys Pro Thr Lys Ile Leu Ser Pro
325 330 335
Glu Tyr Cys Trp Asp Tyr His Ile Gly Leu Pro Ala Asp Ile Lys Leu
340 345 350
Val Lys Met Ser Trp Gln Thr Lys Glu Tyr Asn Val Val Arg Asn Asn
355 360 365
Val

<210> 8
<211> 1131
<212> DNA
<213> *Platyrrhinus helleri*

<220>
<221> CDS
<222> (1)..(1131)
<223>
<400> 8

atg aat gtc aaa gga aaa gta att ctg tcg atg ctg gtt gtc tca act	48
Met Asn Val Lys 5 Lys Val Ile Leu Ser Met Leu Val Val Ser Thr	
1 10 15	
gtg att gtt gtg ttt tgg gaa tat atc aac agc cca gaa ggc tct ttc	96
Val Ile Val Val Phe Trp Glu Tyr Ile Asn Ser Pro Glu Gly Ser Phe	
20 25 30	
ttg tgg ata tat cac tca aag aac cca gaa gtt gat gac agc agt gct	144
Leu Trp Ile Tyr His Ser Lys Asn Pro Glu Val Asp Asp Ser Ser Ala	
35 40 45	
cag aag gac tgg tgg ttt cct ggc tgg ttt aac aat ggg atc cac aat	192
Gln Lys Asp Trp Trp Phe Pro Gly Trp Phe Asn Asn Gly Ile His Asn	
50 55 60	
tat caa caa gag gaa gaa gac aca gac aaa gaa aaa gga aga gag gag	240
Tyr Gln Gln Glu Glu Glu Asp Thr Asp Lys Glu Lys Gly Arg Glu Glu	
65 70 75 80	
gaa caa aaa aag gaa gat gac aca aca gag ctt cgg cta tgg gac tgg	288
Glu Gln Lys Lys 85 Asp Asp Thr Thr 90 Leu Arg Leu Trp Asp Trp	
95	
ttt aat cca aag aaa cgc cca gag gtt atg aca gtg acc caa tgg aag	336
Phe Asn Pro Lys Lys Arg Pro Glu Val Met Thr Val Thr Gln Trp Lys	
100 105 110	
gcg ccg gtt gtg tgg gaa ggc act tac aac aaa gcc atc cta gaa aat	384
Ala Pro Val Val Trp Glu Gly Thr Tyr Asn Lys Ala Ile Leu Glu Asn	
115 120 125	
tat tat gcc aaa cag aaa att acc gtg ggg ttg acg gtt ttt gct att	432
Tyr Tyr Ala Lys Gln Lys Ile Thr Val Gly Leu Thr Val Phe Ala Ile	
130 135 140	
gga aga tat att gag cat tac ttg gag gag ttc gta aca tct gct aat	480
Gly Arg Tyr Ile Glu His Tyr Leu Glu Glu Phe Val Thr Ser Ala Asn	
145 150 155 160	
agg tac ttc atg gtc ggc cac aaa gtc ata ttt tat gtc atg gtg gat	528
Arg Tyr Phe Met Val Gly His Lys Val Ile Phe Tyr Val Met Val Asp	
165 170 175	
gat gtc tcc aag gcg ccg ttt ata gag ctg ggt cct ctg cgt tcc ttc	576
Asp Val Ser Lys Ala Pro Phe Ile Glu Leu Gly Pro Leu Arg Ser Phe	

aaa gtg ttt	180	185	190	624
Lys Val Phe	gag gtc aag cca gag aag agg tgg caa gac atc agc atg	Glu Lys Arg Trp Gln Asp Ile Ser Met		
195	200	205		
atg cgt atg aag acc atc ggg gag cac atc ttg gcc cac atc caa cac	672			
Met Arg Met Lys Thr Ile Gly Glu His Ile Leu Ala His Ile Gln His				
210	215	220		
gag gtt gac ttc ctc ttc tgc atg gat gtg gac cag gtc ttc caa gac	720			
Glu Val Asp Phe Leu Phe Cys Met Asp Val Asp Gln Val Phe Gln Asp				
225	230	235	240	
cat ttt ggg gta gag acc ctg ggc cag tcg gtg gct cag cta cag gcc	768			
His Phe Gly Val Thr Leu Gly Gln Ser Val Ala Gln Leu Gln Ala				
245	250	255		
tggtgg tac aag gca gat cct gat gac ttt acc tat gag agg cgg aaa	816			
Trp Trp Tyr Lys Ala Asp Pro Asp Asp Phe Thr Tyr Glu Arg Arg Lys				
260	265	270		
gag tcg gca gca tat att cca ttt ggc cag ggg gat ttt tat tac cat	864			
Glu Ser Ala Ala Tyr Ile Pro Phe Gly Gln Gly Asp Phe Tyr Tyr His				
275	280	285		
gca gcc att ttt gga gga aca ccg att cag gtt ctc aac atc acc cag	912			
Ala Ala Ile Phe Gly Gly Thr Pro Ile Gln Val Leu Asn Ile Thr Gln				
290	295	300		
gag tgc ttt aag gga atc ctc ctg gac aag aaa aat gac ata gaa gcc	960			
Glu Cys Phe Lys Gly Ile Leu Leu Asp Lys Lys Asn Asp Ile Glu Ala				
305	310	315	320	
gag tgg cat gat gaa agc cac cta aac aag tat ttc ctt ctc aac aaa	1008			
Glu Trp His Asp Glu Ser His Leu Asn Lys Tyr Phe Leu Leu Asn Lys				
325	330	335		
ccc tct aaa atc tta tct cca gaa tac tgc tgg gat tat cat ata ggc	1056			
Pro Ser Lys Ile Leu Ser Pro Glu Tyr Cys Trp Asp Tyr His Ile Gly				
340	345	350		
ctg cct tca gat att aaa act gtc aag cta tca tgg caa aca aaa gag	1104			
Leu Pro Ser Asp Ile Lys Thr Val Lys Leu Ser Trp Gln Thr Lys Glu				
355	360	365		
tat aat ttg gtt aga aag aat gtc tga	1131			
Tyr Asn Leu Val Arg Lys Asn Val				
370	375			

<210> 9
<211> 376
<212> PRT
<213> Platyrrhinus helleri

<400> 9

Met Asn Val Lys Gly Lys Val Ile Leu Ser Met Leu Val Val Ser Thr	1	5	10	15
Val Ile Val Val Phe Trp Glu Tyr Ile Asn Ser Pro Glu Gly Ser Phe	20	25	30	
Leu Trp Ile Tyr His Ser Lys Asn Pro Glu Val Asp Asp Ser Ser Ala	35	40	45	
Gln Lys Asp Trp Trp Phe Pro Gly Trp Phe Asn Asn Gly Ile His Asn	50	55	60	
Tyr Gln Gln Glu Glu Glu Asp Thr Asp Lys Glu Lys Gly Arg Glu Glu	65	70	75	80
Glu Gln Lys Lys Glu Asp Asp Thr Thr Glu Leu Arg Leu Trp Asp Trp	85	90	95	
Phe Asn Pro Lys Lys Arg Pro Glu Val Met Thr Val Thr Gln Trp Lys	100	105	110	

Ala Pro Val Val Trp Glu Gly Thr Tyr Asn Lys Ala Ile Leu Glu Asn
115 120 125

Tyr Tyr Ala Lys Gln Lys Ile Thr Val Gly Leu Thr Val Phe Ala Ile
130 135 140

Gly Arg Tyr Ile Glu His Tyr Leu Glu Glu Phe Val Thr Ser Ala Asn
145 150 155 160

Arg Tyr Phe Met Val Gly His Lys Val Ile Phe Tyr Val Met Val Asp
165 170 175

Asp Val Ser Lys Ala Pro Phe Ile Glu Leu Gly Pro Leu Arg Ser Phe
180 185 190

Lys Val Phe Glu Val Lys Pro Glu Lys Arg Trp Gln Asp Ile Ser Met
195 200 205

Met Arg Met Lys Thr Ile Gly Glu His Ile Leu Ala His Ile Gln His
210 215 220

Glu Val Asp Phe Leu Phe Cys Met Asp Val Asp Gln Val Phe Gln Asp
225 230 235 240

His Phe Gly Val Glu Thr Leu Gly Gln Ser Val Ala Gln Leu Gln Ala
245 250 255

Trp Trp Tyr Lys Ala Asp Pro Asp Asp Phe Thr Tyr Glu Arg Arg Lys
260 265 270

Glu Ser Ala Ala Tyr Ile Pro Phe Gly Gln Gly Asp Phe Tyr Tyr His
275 280 285

Ala Ala Ile Phe Gly Gly Thr Pro Ile Gln Val Leu Asn Ile Thr Gln
290 295 300

Glu Cys Phe Lys Gly Ile Leu Leu Asp Lys Lys Asn Asp Ile Glu Ala
305 310 315 320

Glu Trp His Asp Glu Ser His Leu Asn Lys Tyr Phe Leu Leu Asn Lys
325 330 335

Pro Ser Lys Ile Leu Ser Pro Glu Tyr Cys Trp Asp Tyr His Ile Gly
340 345 350

Leu Pro Ser Asp Ile Lys Thr Val Lys Leu Ser Trp Gln Thr Lys Glu
355 360 365

Tyr Asn Leu Val Arg Lys Asn Val
370 375

<210> 10
<211> 1062
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(1062)
<223>
<400> 10

atg gcc gag gtg ttg cgg acg ctg gcc gga aaa cca aaa tgc cac gca	48
Met Ala Glu Val Leu Arg Thr Leu Ala Gly Lys Pro Lys Cys His Ala	
1 5 10 15	
ctt cga cct atg atc ctt ttc cta ata atg ctt gtc ttg gtc ttg ttt	96
Leu Arg Pro Met Ile Leu Phe Leu Ile Met Leu Val Leu Val Leu Phe	

ggt tac ggg gtc cta agc ccc aga agt cta atg cca gga agc ctg gaa	144
Gly Tyr Gly Val Leu Ser Pro Arg Ser Leu Met Pro Gly Ser Leu Glu	
35	
cgg ggg ttc tgc atg gct gtt agg gaa cct gac cat ctg cag cgc gtc	192
Arg Gly Phe Cys Met Ala Val Arg Glu Pro Asp His Leu Gln Arg Val	
50	
55	
60	
65	
70	
75	
80	
85	
90	
95	
100	
105	
110	
115	
120	
125	
130	
135	
140	
145	
150	
155	
160	
165	
170	
175	
180	
185	
190	
195	
200	
205	
210	
215	
220	
225	
230	
235	
240	
245	
250	
255	
260	
265	
270	
275	
280	
285	
290	
295	
300	
305	
310	
315	
320	
325	
330	
335	
340	
345	
350	
ccg tga	1062
Pro	

<210> 11
<211> 353

<212> PRT
<213> Homo sapiens

<400> 11

```

Met Ala Glu Val Leu Arg Thr Leu Ala Gly Lys Pro Lys Cys His Ala
1      5      10      15
Leu Arg Pro Met Ile Leu Phe Leu Ile Met Leu Val Leu Val Leu Phe
20     25     30
Gly Tyr Gly Val Leu Ser Pro Arg Ser Leu Met Pro Gly Ser Leu Glu
35     40     45
Arg Gly Phe Cys Met Ala Val Arg Glu Pro Asp His Leu Gln Arg Val
50     55     60
Ser Leu Pro Arg Met Val Tyr Pro Gln Pro Lys Val Leu Thr Pro Trp
65     70     75     80
Lys Asp Val Leu Val Val Thr Pro Trp Leu Ala Pro Ile Val Trp Glu
85     90     95
Gly Thr Phe Asn Ile Asp Ile Leu Asn Glu Gln Phe Arg Leu Gln Asn
100    105    110
Thr Thr Ile Gly Leu Thr Val Phe Ala Ile Lys Lys Tyr Val Ala Phe
115    120
Leu Lys Leu Phe Leu Glu Thr Ala Glu Lys His Phe Met Val Gly His
130    135    140
Arg Val His Tyr Tyr Val Phe Thr Asp Gln Leu Ala Ala Val Pro Arg
145    150    155    160
Val Thr Leu Gly Thr Gly Arg Gln Leu Ser Val Leu Glu Val Arg Ala
165    170    175
Tyr Lys Arg Trp Gln Asp Val Ser Met Arg Arg Met Glu Met Ile Ser
180    185    190
Asp Phe Cys Glu Arg Arg Phe Leu Ser Glu Val Asp Tyr Leu Val Cys
195    200    205
Val Asp Val Asp Met Glu Phe Arg Asp His Val Gly Val Glu Ile Leu
210    215    220
Thr Pro Leu Phe Gly Thr Leu His Pro Gly Phe Tyr Gly Ser Ser Arg
225    230    235    240
Glu Ala Phe Thr Tyr Glu Arg Arg Pro Gln Ser Gln Ala Tyr Ile Pro
245    250    255
Lys Asp Glu Gly Asp Phe Tyr Tyr Leu Gly Gly Phe Phe Gly Gly Ser
260    265    270
Val Gln Glu Val Gln Arg Leu Thr Arg Ala Cys His Gln Ala Met Met
275    280    285
Val Asp Gln Ala Asn Gly Ile Glu Ala Val Trp His Asp Glu Ser His
290    295    300
Leu Asn Lys Tyr Leu Leu Arg His Lys Pro Thr Lys Val Leu Ser Pro
305    310    315    320
Glu Tyr Leu Trp Asp Gln Gln Leu Leu Gly Trp Pro Ala Val Leu Arg
325    330    335

```

Lys Leu Arg Phe Thr Ala Val Pro Lys Asn His Gln Ala Val Arg Asn
340 345 350

Pro

<210> 12
<211> 1065
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(1065)
<223>
<400> 12

atg gcc gag gtg ttg cgg acg ctg gcc gga aaa cca aaa tgc cac gca	48
Met Ala Glu Val 5 Arg Thr Leu Ala Gly Lys Pro Lys Cys His Ala	
1 10 15	
ctt cga cct atg atc ctt ttc cta ata atg ctt gtc ttg gtc ttg ttt	96
Leu Arg Pro Met Ile Leu Phe Leu Ile Met Leu Val Leu Val Leu Phe	
20 25 30	
ggt tac ggg gtc cta agc ccc aga agt cta atg cca gga agc ctg gaa	144
Gly Tyr Gly Val Leu Ser Pro Arg Ser Leu Met Pro Gly Ser Leu Glu	
35 40 45	
cgg ggg ttc tgc atg gct gtt agg gaa cct gac cat ctg cag cgc gtc	192
Arg Gly Phe Cys Met Ala Val Arg Glu Pro Asp His Leu Gln Arg Val	
50 55 60	
tcg ttg cca agg atg gtc tac ccc cag cca aag gtg ctg aca ccg tgt	240
Ser Leu Pro Arg Met Val Tyr Pro Gln Pro Lys Val Leu Thr Pro Cys	
65 70 75 80	
agg aag gat gtc ctc gtg gtg acc cct tgg ctg gct ccc att gtc tgg	288
Arg Lys Asp Val 85 Val Val Thr Pro Leu Ala Pro Ile Val Trp	
90 95	
gag ggc acg ttc aac atc gac atc ctc aac gag cag ttc agg ctc cag	336
Glu Gly Thr Phe Asn Ile Asp Ile Leu Asn Glu Gln Phe Arg Leu Gln	
100 105 110	
aac acc acc att ggg tta act gtg ttt gcc atc aag aaa tac gtg gct	384
Asn Thr Thr Ile Gly Leu Thr Val Phe Ala Ile Lys Lys Tyr Val Ala	
115 120 125	
ttc ctg aag ctg ttc ctg gag acg gcg gag aag cac ttc atg gtg ggc	432
Phe Leu Lys Leu Phe Leu Glu Thr Ala Glu Lys His Phe Met Val Gly	
130 135 140	
cac cgt gtc cac tac tat gtc ttc acc gac cag ccg gcc gcg gtg ccc	480
His Arg Val His Tyr Tyr Val Phe Thr Asp Gln Pro Ala Ala Val Pro	
145 150 155 160	
cgc gtg acg ctg ggg acc ggt cgg cag ctg tca gtg ctg gag gtg ggc	528
Arg Val Thr Leu Gly Thr Gly Arg Gln Leu Ser Val Leu Glu Val Gly	
165 170 175	
gcc tac aag cgc tgg cag gac gtg tcc atg cgc cgc atg gag atg atc	576
Ala Tyr Lys Arg Trp Gln Asp Val Ser Met Arg Arg Met Glu Met Ile	
180 185 190	
agt gac ttc tgc gag cgg cgc ttc ctc agc gag gtg gat tac ctg gtg	624
Ser Asp Phe Cys Glu Arg Arg Phe Leu Ser Glu Val Asp Tyr Leu Val	
195 200 205	
tgc gtg gac gtg gac atg gag ttc cgc gac cat gtg ggc gtg gag atc	672
Cys Val Asp Val Asp Met Glu Phe Arg Asp His Val Gly Val Glu Ile	
210 215 220	
ctg act ccg ctg ttc ggc acc ctg cac ccc agc ttc tac gga agc agc	720
Leu Thr Pro Leu Phe Gly Thr Leu His Pro Ser Phe Tyr Gly Ser Ser	
225 230 235 240	
cgg gag gcc ttc acc tac gag cgc cgg ccc cag tcc cag gcc tac atc	768
Arg Glu Ala Phe Thr Tyr Glu Arg Arg Pro Gln Ser Gln Ala Tyr Ile	
245 250 255	
ccc aag gac gag ggc gat ttc tac tac atg ggg gcg ttc ttc ggg ggg	816

Pro	Lys	Asp	Glu	Gly	Asp	Phe	Tyr	Tyr	Met	Gly	Ala	Phe	Phe	Gly	Gly		
			260				265						270				
tcg	gtg	caa	gag	gtg	cag	cgg	ctc	acc	agg	gcc	tgc	cac	cag	gcc	atg	864	
Ser	Val	Gln	Glu	Val	Gln	Arg	Leu	Thr	Arg	Ala	Cys	His	Gln	Ala	Met		
		275					280					285					
atg	gtc	gac	cag	gcc	aac	ggc	atc	gag	gcc	gtg	tgg	cac	gac	gag	agc	912	
Met	Val	Asp	Gln	Ala	Asn	Gly	Ile	Glu	Ala	Val	Trp	His	Asp	Glu	Ser		
		290				295					300						
cac	ctg	aac	aag	tac	cta	ctg	cgc	cac	aaa	ccc	acc	aag	gtg	ctc	tcc	960	
His	Leu	Asn	Lys	Tyr	Leu	Leu	Arg	His	Lys	Pro	Thr	Lys	Val	Leu	Ser		
305					310				315					320			
ccc	gag	tac	ttg	tgg	gac	cag	cag	ctg	ctg	ggc	tgg	ccc	gcc	gtc	ctg	1008	
Pro	Glu	Tyr	Leu	Trp	Asp	Gln	Gln	Leu	Leu	Gly	Trp	Pro	Ala	Val	Leu		
			325					330					335				
agg	aag	ctg	agg	ttc	act	gcg	gtg	ccc	aag	aac	cac	cag	gcg	gtc	cgg	1056	
Arg	Lys	Leu	Arg	Phe	Thr	Ala	Val	Pro	Lys	Asn	His	Gln	Ala	Val	Arg		
			340					345					350				
aac	ccg	tga														1065	
Asn	Pro																

<210> 13
<211> 354
<212> PRT
<213> Homo sapiens

<400> 13

Met	Ala	Glu	Val	Leu	Arg	Thr	Leu	Ala	Gly	Lys	Pro	Lys	Cys	His	Ala
1				5					10					15	
Leu	Arg	Pro	Met	Ile	Leu	Phe	Leu	Ile	Met	Leu	Val	Leu	Val	Leu	Phe
		20					25					30			
Gly	Tyr	Gly	Val	Leu	Ser	Pro	Arg	Ser	Leu	Met	Pro	Gly	Ser	Leu	Glu
		35					40					45			
Arg	Gly	Phe	Cys	Met	Ala	Val	Arg	Glu	Pro	Asp	His	Leu	Gln	Arg	Val
	50				55					60					
Ser	Leu	Pro	Arg	Met	Val	Tyr	Pro	Gln	Pro	Lys	Val	Leu	Thr	Pro	Cys
65				70					75					80	
Arg	Lys	Asp	Val	Leu	Val	Val	Thr	Pro	Trp	Leu	Ala	Pro	Ile	Val	Trp
			85					90					95		
Glu	Gly	Thr	Phe	Asn	Ile	Asp	Ile	Leu	Asn	Glu	Gln	Phe	Arg	Leu	Gln
		100					105					110			
Asn	Thr	Thr	Ile	Gly	Leu	Thr	Val	Phe	Ala	Ile	Lys	Lys	Tyr	Val	Ala
		115					120					125			
Phe	Leu	Lys	Leu	Phe	Leu	Glu	Thr	Ala	Glu	Lys	His	Phe	Met	Val	Gly
	130				135						140				
His	Arg	Val	His	Tyr	Tyr	Val	Phe	Thr	Asp	Gln	Pro	Ala	Ala	Val	Pro
145					150				155					160	
Arg	Val	Thr	Leu	Gly	Thr	Gly	Arg	Gln	Leu	Ser	Val	Leu	Glu	Val	Gly
			165					170					175		
Ala	Tyr	Lys	Arg	Trp	Gln	Asp	Val	Ser	Met	Arg	Arg	Met	Glu	Met	Ile
		180					185						190		
Ser	Asp	Phe	Cys	Glu	Arg	Arg	Phe	Leu	Ser	Glu	Val	Asp	Tyr	Leu	Val
		195					200					205			

Cys Val Asp Val Asp Met Glu Phe Arg Asp His Val Gly Val Glu Ile
210 215 220

Leu Thr Pro Leu Phe Gly Thr Leu His Pro Ser Phe Tyr Gly Ser Ser
225 230 235 240

Arg Glu Ala Phe Thr Tyr Glu Arg Arg Pro Gln Ser Gln Ala Tyr Ile
245 250 255

Pro Lys Asp Glu Gly Asp Phe Tyr Tyr Met Gly Ala Phe Phe Gly Gly
260 265 270

Ser Val Gln Glu Val Gln Arg Leu Thr Arg Ala Cys His Gln Ala Met
275 280 285

Met Val Asp Gln Ala Asn Gly Ile Glu Ala Val Trp His Asp Glu Ser
290 295 300

His Leu Asn Lys Tyr Leu Leu Arg His Lys Pro Thr Lys Val Leu Ser
305 310 315 320

Pro Glu Tyr Leu Trp Asp Gln Gln Leu Leu Gly Trp Pro Ala Val Leu
325 330 335

Arg Lys Leu Arg Phe Thr Ala Val Pro Lys Asn His Gln Ala Val Arg
340 345 350

Asn Pro

<210> 14
<211> 23
<212> DNA
<213> Artificial

<400> 14

ggcctgtact acatttgctt gga

23

<210> 15
<211> 26
<212> DNA
<213> Artificial

<400> 15

gaaatagtgt caagtttcca tcacaa

26

<210> 16
<211> 55
<212> DNA
<213> Artificial

<400> 16

cgatgtggct gcggagccac cggcaggttaa tcctgttgat gctgattgtc tcaac

55

<210> 17
<211> 50
<212> PRT
<213> Mus musculus

<400> 17

Met Ile Thr Met Leu Gln Asp Leu His Val Asn Lys Ile Ser Met Ser
1 5 10 15

Arg Ser Lys Ser Glu Thr Ser Leu Pro Ser Ser Arg Ser Gly Ser Gln
20 25 30

Glu Lys Ile Met Asn Val Lys Gly Lys Val Ile Leu Leu Met Leu Ile
35 40 45

Val Ser
50

<210> 18
<211> 17
<212> PRT
<213> Artificial

<400> 18

Met Trp Leu Arg Ser His Arg Gln Val Ile Leu Leu Met Leu Ile Val
1 5 10 15

Ser

<210> 19
<211> 17
<212> PRT
<213> Artificial

<400> 19

Met Trp Leu Arg Ser His Arg Gln Val Val Leu Ser Met Leu Leu Val
1 5 10 15

Ser

<210> 20
<211> 15
<212> PRT
<213> Sus scrofa

<400> 20

Met Asn Val Lys Gly Arg Val Val Leu Ser Met Leu Leu Val Ser
1 5 10 15